JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 4, NUMBER 8 AUGUST 1963

Expansions Associated with a Pair of Singular First-Order Differential Equations

B. W. ROOS AND W. C. SANGREN*
John Jay Hopkins Laboratory for Pure and Applied Science, General Atomic Division of General Dynamics Corporation, San Diego, California
(Received 31 October 1962)

This paper reports on the continuation of a study of the spectral properties of a Dirac operator. The analytical methods developed by Weyl and Titchmarsh for the analysis of the Sturm–Liouville equation are extended to the investigation of a system of two singular first-order differential equations. Expansions associated with the system are established and a convergence theorem is presented.

1. INTRODUCTION

THE analytical methods developed by H. Weyl[1] and E. C. Titchmarsh[2] for the solution of the Sturm–Liouville equation,

$$x''(r) + [\lambda - V(r)]x(r) = 0 \quad (' \equiv d/dr) \tag{1}$$

are powerful tools for the investigation of the singular cases of the second-order differential equations of mathematical physics. For example, since Eq. (1) corresponds to the radial wave equation of a nonrelativistic particle in a central field, the methods referred to may be advantageously applied to determine the properties of the Schrödinger operator for singular potentials.[3]

The relativistic counterpart of Eq. (1), i.e., the Dirac radial relativistic wave equation for a particle in a central field, takes the form of a system of two first-order differential equations:

$$\begin{aligned} x_1'(r) - [\lambda a(r) + b(r)]x_2(r) &= 0, \\ x_2''(r) + [\lambda c(r) + d(r)]x_1(r) &= 0. \end{aligned} \tag{2}$$

It is the purpose of this paper to generalize the methods of Weyl and Titchmarsh in order to investigate the expansions associated with this system of Eqs. (2). The present paper is a continuation of another paper by the authors on the spectrum of the Dirac radial wave equations.[4]

The arguments here are presented for the case of a semi-infinite interval $(0, \infty)$ where the point $r = 0$ is assumed to be a regular point and the system is only singular for $r \to \infty$. The coefficients a, b, c, d are assumed to be real-valued continuous functions of r, and $a(r)$ and $b(r)$ are furthermore assumed to be positive on any finite interval. The boundary condition considered at $r = 0$ is given by

$$k^{-1}(0) \cos \beta x_1(0) + k(0) \sin \beta x_2(0) = 0. \tag{3}$$

where β is a real constant and where $k(r) = [a(r)/c(r)]^{\frac{1}{2}}$.

2. PRELIMINARIES

On the finite interval $(0, r^0)$, let $v(r, \lambda) = [v_1(r, \lambda), v_2(r, \lambda)]$ and $w(r, \lambda) = [w_1(r, \lambda), w_2(r, \lambda)]$ be two vector solutions of system (2) that satisfy the conditions

* Present address: Computer Applications, Inc., San Diego 10, California.

[1] H. Weyl, Ann. Math. **68**, 220 (1910).

[2] E. C. Titchmarsh, *Eigenfunction Expansions* (Clarendon Press, Oxford, England, 1946); Proc. Lond. Math. Soc. **11**, 159 (1961).

[3] K. Case, Phys. Rev. **80**, 797 (1950).

[4] B. W. Roos and W. C. Sangren, J. Math. Phys. **3**, 882 (1962).

$$v_1(0) = -\sin \beta k(0), \qquad v_2(0) = \cos \beta k^{-1}(0),$$
$$w_1(0) = -\cos \beta k(0), \qquad w_2(0) = -\sin \beta k^{-1}(0).$$

The Wronskian of v and w is defined by

$$W_r(v, w) = v_1(r)w_2(r) - v_2(r)w_1(r).$$

Since $W_r(v, w)$ is independent of r and $W_0(v, w) = 1$, $W_r(v, w) = 1$ and v and w are linearly independent solutions. A solution of (2) may be written as $w(r, \lambda) + l(\lambda)v(r, \lambda)$, and if this solution satisfies Sturmian boundary conditions at a point $r = r_1$, the eigenvalues will be real, nondegenerate, and discrete, and extend from $\lambda = -\infty$ to $\lambda = \infty$. The corresponding eigenfunctions are real functions of r. For the singular case, the spectrum can be investigated by taking the limit of the general solution as $r^0 \to \infty$. As in the case of singular second-order differential equations, it can be shown[4] by a limit-point, limit-circle argument that for $\operatorname{Im} \lambda \neq 0$, the system (2) will have a vector solution

$$z(r, \lambda) = w(r, \lambda) + m(\lambda)v(r, \lambda) \tag{4}$$

belonging to the class of square-integrable functions $L^2(r, \infty)$. The function $m(\lambda)$ depends upon the limit of circles in the complex λ plane, and for $r^0 \to \infty$ is either a limit point or a point on a limit circle. In the limit-circle case, all solutions are in the class $L^2(r_0, \infty)$. Furthermore, $m(\lambda)$ is analytic for $\operatorname{Im} \lambda \neq 0$ and $m(\bar{\lambda}) = \overline{m(\lambda)}$.

3. GENERAL TRANSFORMATION OF THE BASIC EQUATIONS

The asymptotic behavior of the solutions of Eqs. (2) for large values of the parameter λ and for large values of one or more of the coefficients when the independent variable approaches a singular point, can be conveniently investigated by using the following transformation. The independent variable r is replaced by

$$\alpha(r) = \int_0^r \{[\lambda a(s) + b(s)][\lambda c(s) + d(s)]\}^{\frac{1}{2}}\, ds. \tag{5}$$

Formally, this transformation gives

$$\begin{aligned} du_1/d\alpha &= u_2, \\ du_2/d\alpha &= -u_1 + R(\lambda, a, b, c, d)u_1, \end{aligned} \tag{6}$$

where the components of the vector $u(r)$ are defined by

$$\begin{aligned} u_1(r) &= F(r)x_2(r), \\ u_2(r) &= -F^{-1}(r)x_1(r) + G(r)x_2(r), \end{aligned} \tag{7}$$

and

$$\begin{aligned} F(r, \lambda) &= [\lambda a(r) + b(r)]^{\frac{1}{2}}[\lambda c(r) + d(r)]^{-\frac{1}{2}}, \\ G(r, \lambda) &= (\alpha')^{-1}F'(r, \lambda), \\ R(\lambda, a, b, c, d) &= \alpha(')^{-1}G'(r, \lambda)F^{-1}(r, \lambda). \end{aligned} \tag{8}$$

The transformed equations (6) have a form similar to (2). However, the coefficients on the right-hand side of Eqs. (6) do not in general become large when λ does or when one or more of the coefficients a, b, c, d, do.

It can be verified directly that a solution of (6) satisfies the integral equations

$$\begin{aligned} U_1(r, \lambda) &= U_2(0) \sin \alpha(r) + U_1(0) \cos \alpha(r) \\ &\quad + \int_0^r U_1(s)S(s) \sin [\alpha(r) - \alpha(s)]\, ds, \\ U_2(r, \lambda) &= -U_1(0) \sin \alpha(r) + U_2(0) \cos \alpha(r) \\ &\quad + \int_0^r U_1(s)S(s) \cos [\alpha(r) - \alpha(s)]\, ds, \end{aligned} \tag{9}$$

where

$$S(r) = G'(r)F^{-1}(r). \tag{10}$$

4. ORDER PROPERTIES FOR LARGE VALUES OF λ

It will be assumed that a and c are bounded away from 0 and ∞ for all finite r. For a fixed finite r and large $|\lambda|$, we obtain from (5),

$$\alpha(r) = \lambda h(r) + g(r) + O(\lambda^{-1}), \tag{11}$$

where

$$h(r) = \int_0^r [a(s)c(s)]^{\frac{1}{2}}\, ds$$

and (12)

$$g(r) = \frac{1}{2} \int_0^r \frac{b(s)c(s) + a(s)d(s)}{[a(s)c(s)]^{\frac{1}{2}}}\, ds.$$

It is not difficult to verify also that for large $|\lambda|$,

$$\begin{aligned} F(r) &= [a(r)/c(r)]^{\frac{1}{2}} + O(\lambda^{-1}) = k(r) + O(\lambda^{-1}), \\ F^{-1}(r) &= [c(r)/a(r)]^{\frac{1}{2}} + O(\lambda^{-1}) = k^{-1}(r) + O(\lambda^{-1}), \\ G(r) &= O(\lambda^{-1}), \qquad S(r) = O(\lambda^{-1}). \end{aligned} \tag{13}$$

Next, consider the order properties of $U_1(0)$ and $U_2(0)$ for large $|\lambda|$. If we assume, without loss of generality, that the boundary conditions (3) are satisfied for $x_1(0) = -\sin \beta k(0)$ and $x_2(0) = \cos \beta k^{-1}(0)$, it follows directly from the transformation that

$$\begin{aligned} U_1(0) &= \cos \beta + O(\lambda^{-1}), \\ U_2(0) &= -\sin \beta + O(\lambda^{-1}). \end{aligned} \tag{14}$$

Now let $\lambda = \sigma + i\tau$, where $\tau > 0$, and furthermore let

$$U_1(r, \lambda) = H_1(r, \lambda)e^{\tau h(r)}, \quad U_2(r, \lambda) = H_2(r, \lambda)e^{\tau h(r)}. \tag{15}$$

It follows from Eq. (9) that

$$H_1(r, \lambda) = [U_2(0) \sin \alpha(r) + U_1(0) \cos \alpha(r)]e^{-\tau h(r)} + \int_0^r e^{-\tau[h(r)-h(s)]}H_1(s)S(s) \sin [\alpha(r) - \alpha(s)]\, ds,$$
$$H_2(r, \lambda) = [-U_1(0) \sin \alpha(r) + U_2(0) \cos \alpha(r)]e^{-\tau h(r)} + \int_0^r e^{-\tau[h(r)-h(s)]}H_1(s)S(s) \cos [\alpha(r) - \alpha(s)]\, ds. \tag{16}$$

If

$$\mu = \max_{s \geq r} [|H_1|, |H_2|],$$

these equations yield for large $|\lambda|$, after absolute values are taken,

$$\mu < |U_1(0)| + |U_2(0)| + \mu \int_0^r |S(s)|\, ds.$$

It now follows from a lemma of a previous paper[5] that

$$\mu \leq [|\cos \beta| + |\sin \beta|] \exp\left[\int_0^r |S(s)|\, ds\right],$$

provided that $\int_0^r |S(s)|\, ds$ exists. $H_1(r, \lambda)$ and $H_2(r, \lambda)$ are therefore seen to be bounded for all r, provided that $\int_0^\infty |S(s, \lambda)|\, ds$ is uniformly convergent with respect to $\lambda > \rho \gg 0$. It follows immediately that

$$U_1(r, \lambda) = O\{e^{\tau h(r)}\} \text{ and } U_2(r, \lambda) = O\{e^{\tau h(r)}\}. \tag{17}$$

In the light of later developments, it seemed natural to impose the more stringent condition that for large $|\lambda|$,

$$\int_0^\infty |S(s, \lambda)|\, ds = O(|\lambda|^{-1}). \tag{18}$$

From the order properties for $U_1(r, \lambda)$ and $U_2(r, \lambda)$, and the conditions imposed on the integral $\int_0^\infty |S(s, \lambda)|\, ds$, we obtain from the integral equations (9) the following relations for large values of $|\lambda|$:

$$U_1(r, \lambda) = \cos [\alpha(r) + \beta] + O\{e^{\tau h(r)} |\lambda|^{-1}\},$$
$$U_2(r, \lambda) = -\sin [\alpha(r) + \beta] + O\{e^{\tau h(r)} |\lambda|^{-1}\}. \tag{19}$$

[5] B. W. Roos and W. C. Sangren, Proc. Am. Math. Soc. 12, 468 (1961).

Further, from the transformation,

$$x_1(r, \lambda) = -k(r) \sin [\alpha(r) + \beta] + O\{e^{\tau h(r)} |\lambda|^{-1}\},$$
$$x_2(r, \lambda) = +k^{-1}(r) \cos [\alpha(r) + \beta] + O\{e^{\tau h(r)} |\lambda|^{-1}\}. \tag{20}$$

5. A SPECIAL SOLUTION: $X(r, \lambda)$

Before proving a general convergence theorem, it is desirable to obtain a special solution of Eqs. (2) that is small for large r when the imaginary part of λ is large and positive. Again, it is more convenient to consider the solutions u_1 and u_2 of the transformed equations (6). Actually, we first consider the integral equations

$$U_1(r, \lambda) = e^{i\alpha(r)} - \frac{i}{2}\int_0^r e^{i[\alpha(r)-\alpha(s)]}S(s)U_1(s)\, ds - \frac{i}{2}\int_r^\infty e^{i[\alpha(s)-\alpha(r)]}S(s)U_1(s)\, ds,$$
$$U_2(r, \lambda) = ie^{i\alpha(r)} + \frac{1}{2}\int_0^r e^{i[\alpha(r)-\alpha(s)]}S(s)U_1(s)\, ds - \frac{1}{2}\int_r^\infty e^{i[\alpha(s)-\alpha(r)]}S(s)U_1(s)\, ds. \tag{21}$$

It is not difficult to verify that the solutions of these integral equations satisfy the differential equations (6).

In the study of the spectra for Eqs. (2), it was shown[5] that for the various cases of interest, either Im $\alpha(r) \to +\infty$ or Im $\alpha(r) \to -\infty$ as $r \to \infty$ and $\tau > 0$. In order to avoid duplication and a clumsy notation, attention is here kept fixed on the case Im $\alpha(r) \to +\infty$. The following arguments are easily duplicated for the case Im $\alpha(r) \to -\infty$.

A special solution of Eqs. (6) can be obtained by successive substitution as follows: First, let

$$U_{11}(r, \lambda) = e^{i\alpha(r)}, \quad U_{21}(r, \lambda) = ie^{i\alpha(r)}. \tag{22}$$

Next, for $n \geq 1$, let

$$U_{1,n+1}(r, \lambda) = e^{i\alpha(r)} - \frac{i}{2}\int_0^r e^{i[\alpha(r)-\alpha(s)]}S(s)U_{1,n}(s)\, ds - \frac{i}{2}\int_r^\infty e^{i[\alpha(s)-\alpha(r)]}S(s)U_{1,n}(s)\, ds,$$
$$U_{2,n+1}(r, \lambda) = ie^{i\alpha(r)} + \frac{1}{2}\int_0^r e^{i[\alpha(r)-\alpha(s)]}S(s)U_{1,n}(s)\, ds - \frac{1}{2}\int_r^\infty e^{i[\alpha(s)-\alpha(r)]}S(s)U_{1,n}(s)\, ds. \tag{23}$$

Then,

$$U_{1,2} - U_{1,1} = -\frac{i}{2}e^{i\alpha(r)}\left[\int_0^r S(s)\,ds + \int_r^\infty e^{2i[\alpha(s)-\alpha(r)]}S(s)\,ds\right],$$

and

$$|U_{1,2} - U_{1,1}| \leq \frac{e^{-\mathrm{Im}\alpha(r)}}{2}\int_0^\infty |S(s)|\,ds \leq \frac{Ae^{-\mathrm{Im}\alpha(r)}}{|\lambda|},$$

because

$$\int_0^\infty |S(s)|\,ds = O(\lambda^{-1}).$$

Similarly, it can be shown that

$$|U_{1,n+1} - U_{1,n}| \leq [A/|\lambda|]^n e^{-\mathrm{Im}\alpha(r)},$$
$$|U_{2,n+1} - U_{2,n}| \leq [A/|\lambda|]^n e^{-\mathrm{Im}\alpha(r)}.$$

Hence, the series $\sum_{n=1}^\infty [U_{1,n+1}(r) - U_{1,n}(r)]$ and the series $\sum_{n=1}^\infty [U_{2,n+1}(r) - U_{2,n}(r)]$ are convergent if $|\lambda| > A$. Therefore, by definition, let

$$U_1(r) = \lim_{n\to\infty} U_{1,n}(r) \quad \text{and} \quad U_2(r) = \lim_{n\to\infty} U_{2,n}(r).$$

Now for every n,

$$|U_{1,n}(x)| \leq |U_{1,1}| + |U_{1,2} - U_{1,1}| + \cdots + |U_{1,n} - U_{1,n-1}| \leq e^{-\mathrm{Im}\alpha(r)}\left[\frac{1 - (A/|\lambda|)^n}{1 - A/|\lambda|}\right],$$

and for $n \to \infty$ and $|\lambda| > A$,

$$|U_1(r)| = \lim_{n\to\infty} |U_{1,n}(x)| \leq e^{-\mathrm{Im}\alpha(r)}/(1 - A/|\lambda|).$$

Similarly,

$$|U_2(r)| \leq \lim_{n\to\infty} |U_{2,n}(r)| \leq e^{-\mathrm{Im}\alpha(r)}/(1 - A/|\lambda|).$$

By dominated convergence, it follows that the limit operations may be taken under the integral signs and that $U_1(r)$ and $U_2(r)$ satisfy Eqs. (21) and consequently, Eqs. (6).

Consider Eqs. (21) and the solution $U = [U_1, U_2]$. It is easily seen from these equations that, as $r \to \infty$,

$$U_1(r, \lambda) = e^{i\alpha(r)}\{K_1(\lambda) + o(1)\},$$
$$U_2(r, \lambda) = e^{i\alpha(r)}\{K_2(\lambda) + o(1)\},$$

where

$$K_1(\lambda) = 1 - \frac{i}{2}\int_0^\infty e^{-i\alpha(s)}S(s)U_1(s)\,ds,$$

$$K_2(\lambda) = i + \frac{1}{2}\int_0^\infty e^{-i\alpha(s)}S(s)U_1(s)\,ds,$$

and

$$K_2(\lambda) = iK_1(\lambda).$$

By use of the general transformation, these special solutions can be converted to solutions of Eqs. (2). We have

$$X_1(r, \lambda) = F(r)\{G(r)U_1(r) - U_2(r)\},$$
$$X_2(r, \lambda) = F^{-1}(r)U_1(r).$$

By substitution for U_1 and U_2, we have

$$X_1(r, \lambda) = F(r)e^{i\alpha(r)}\{G(r)K_1(\lambda) - K_2(\lambda) + o(1)\},$$
$$X_2(r, \lambda) = F^{-1}(r)e^{i\alpha(r)}\{K_1(\lambda) + o(1)\}. \tag{24}$$

6. ORDER PROPERTIES OF $X(r, \lambda)$

Let $v(r, \lambda)$ be the vector solution mentioned in Sec. 2. The order properties of $v_1(r, \lambda)$ and $v_2(r, \lambda)$ for large λ are given by Eqs. (20):

$$v_1(r, \lambda) = -k(r)\sin[\alpha(r) + \beta] + O\{e^{\tau h(r)}|\lambda|^{-1}\},$$
$$v_2(r, \lambda) = +k^{-1}(r)\cos[\alpha(r) + \beta] + O\{e^{\tau h(r)}|\lambda|^{-1}\}.$$

Now $v(r, \lambda)$ and $X(r, \lambda)$ are linearly independent solutions, because $X(r, \lambda)$ becomes small for large r whereas $v(r, \lambda)$ becomes large for large r. Next take $X(r, \lambda)$ to be proportional to $z(r, \lambda)$, that is,

$$z(r, \lambda) = L(\lambda)X(r, \lambda).$$

This is a valid assumption if $X(r, \lambda)$ belongs to $L^2(0, \infty)$. That this is true may be verified by considering formulas (22). For large $|\lambda|$ and r, these formulas reduce to

$$X_1(r, \lambda) = -k(r)e^{i[\lambda h(r) + g(r)]}K_2(\lambda),$$
$$X_2(r, \lambda) = k^{-1}(r)e^{i[\lambda h(r) + g(r)]}K_1(\lambda),$$

and these last two functions are in $L^2(0, \infty)$.

The order properties for $w(r, \lambda)$ and $v(r, \lambda)$ for large r can be derived directly from Eqs. (7). They have been derived previously[4] and are listed here for convenient reference. For $e^{-i\alpha(r)} \to \infty$,

$$U_1(r, \lambda) = e^{-i\alpha(r)}\{M(\lambda) + o(1)\},$$
$$U_2(r, \lambda) = e^{-i\alpha(r)}\{N(\lambda) + o(1)\},$$

where

$$M(\lambda) = -\frac{U_2(0)}{2i} + \frac{U_1(0)}{2} - \frac{1}{2i}\int_0^\infty e^{i\alpha(s)}S(s)U_1(s)\,ds,$$

$$N(\lambda) = \frac{U_1(0)}{2i} + \frac{U_2(0)}{2} + \frac{1}{2}\int_0^\infty e^{i\alpha(s)}S(s)U_1(s)\,ds.$$

From the basic transformation, it follows that

$$\begin{aligned} v_1(r, \lambda) &= -F(r)U_2(r, \lambda) + F(r)G(r)U_1(r, \lambda) \\ &= F(r)e^{-i\alpha(r)}\{-N(\lambda) + G(r)M(\lambda) + o(1)\}, \\ v_2(r, \lambda) &= F^{-1}(r)U_1(r, \lambda) \\ &= F^{-1}(r)e^{-i\alpha(r)}\{M(\lambda) + o(1)\}. \end{aligned}$$

Consider next the Wronskian

$$\begin{aligned} W_r(v, z) &= W_r(v, LX) \\ &= L(\lambda)[v_1X_2 - v_2X_1] \\ &= 2iL(\lambda)K_1(\lambda)M(\lambda) + o(1). \end{aligned}$$

Because $W_r(v, w) = 1$, it follows that

$$L(\lambda) = [2iK_1(\lambda)M(\lambda)]^{-1}.$$

Hence,

$$z(r, \lambda) = X(r, \lambda)[2iK_1(\lambda)M(\lambda)]^{-1},$$

and therefore

$$|z(r, \lambda)| = |X(r, \lambda)|\ \{2\ |K_1(\lambda)|\ |M(\lambda)|\}^{-1}.$$

The previously obtained bounds for $U_1(r, \lambda)$ and $U_2(r, \lambda)$ and the basic transformations therefore give

$$\begin{aligned} |z_1(r, \lambda)| &\le \{2\ |K_1(\lambda)|\ |M(\lambda)|\}^{-1}\{1 - A/|\lambda|\}^{-1} \\ &\quad \times e^{-\mathrm{Im}\alpha(r)}\{|F(r)| + |F(r)|\ |G(r)|\}, \\ |z_2(r, \lambda)| &\le \{2\ |K_1(\lambda)|\ |M(\lambda)|\}^{-1}\{1 - A/|\lambda|\}^{-1} \\ &\quad \times e^{-\mathrm{Im}\alpha(r)}\ |F^{-1}(r)|. \end{aligned}$$

Now M and K both exist and are bounded away from zero, because (as was shown) $W_r(v, z) = 1 = 2iMK_1L$. Finally, we can conclude that for large λ,

$$|z(r, \lambda)| \le CE(r)e^{-\mathrm{Im}\alpha(r)},$$

where C is a constant and

$$E(r) = \max\ [|F^{-1}(r)|,\ |F(r)|\ |G(r)| + |F(r)|].$$

For later reference, we need an expression for $z(r, \lambda)$ which holds for all r and for large λ. Because

$$z(r, \lambda) = X(r, \lambda)\{2iK_1(\lambda)M(\lambda)\},$$

and $X(r, \lambda)$ in turn depends only upon the special solution $U_1(r, \lambda)$ and $U_2(r, \lambda)$, it is desirable to consider Eqs. (21). These equations can be written in the form

$$\begin{aligned} U_1(r, \lambda) &= e^{i\alpha(r)}\Big\{1 - \frac{i}{2}\int_0^\infty e^{-i\alpha(s)}S(s)U_1(s)\ ds \\ &\quad + \frac{i}{2}\int_r^\infty e^{i\alpha(s)}S(s)U_1(s)\ ds \\ &\quad - \frac{i}{2}\int_r^\infty e^{i[\alpha(s)-\alpha(r)]}S(s)U_1(s)\ ds\Big\} \\ &= e^{i\alpha(r)}\Big\{K_1(\lambda) \\ &\quad + \frac{i}{2}\int_r^\infty e^{-i\alpha(s)}S(s)U_1(s)\ ds\Big\}[1 + O(\lambda^{-1})], \\ U_2(r, \lambda) &= e^{i\alpha(r)}\Big\{iK_1(\lambda) \\ &\quad - \frac{1}{2}\int_r^\infty e^{-i\alpha(s)}S(s)U_1(s)\ ds\Big\}[1 + O(\lambda^{-1})]. \end{aligned}$$

It is apparent that these relations may be written in the form

$$\begin{aligned} U_1(r, \lambda) &= C_1(r, \lambda)e^{+i\alpha(r)}[1 + O(\lambda^{-1})], \\ U_2(r, \lambda) &= C_2(r, \lambda)e^{+i\alpha(r)}[1 + O(\lambda^{-1})], \end{aligned}$$

where

$$C_2(r, \lambda) = iC_1(r, \lambda),$$

and

$$C_1(r, \lambda) = K_1(\lambda) + \frac{i}{2}\int_r^\infty C_1(s, \lambda)S(s)\ ds.$$

It is easily verified that this integral equation has the solution

$$C_1(r, \lambda) = D(\lambda)\ \exp\left[\frac{i}{2}\int_0^\infty S(s)\ ds\right],$$

where

$$D(\lambda) = K_1(\lambda)\ \exp\left[\frac{i}{2}\int_0^\infty S(s)\ ds\right].$$

Hence,

$$C_1(r, \lambda) = K_1(\lambda)\ \exp\left[\int_r^\infty S(s)\ ds\right].$$

It now follows directly that

$$\begin{aligned} X_2(r, \lambda) &= F^{-1}(r)C_1(r, \lambda)e^{i\alpha(r)}[1 + O(|\lambda|^{-1})], \\ X_1(r, \lambda) &= [F(r)G(r)C_1(r, \lambda) \\ &\quad - F(r)C_2(r, \lambda)]e^{i\alpha(r)}[1 + O(|\lambda|^{-1})], \end{aligned}$$

and therefore that

$$\begin{aligned} z_1(r, \lambda) &= -\frac{1}{2}k(r)M^{-1}(\lambda) \\ &\quad \times \exp\left[\int_r^\infty S(s)\ ds + i\alpha(r)\right][1 + O(\lambda^{-1})], \end{aligned}$$

$$z_2(r,\lambda) = \frac{1}{2i}k^{-1}(r)M^{-1}(\lambda) \times \exp\left[\int_r^\infty S(s)\,ds + i\alpha(r)\right][1 + O(\lambda^{-1})].$$

7. A CONVERGENCE THEOREM

Theorem. If $f(r) = [f_1(r), f_2(r)]$ *is such that the integral*

$$\int_0^\infty E(s)\{c(s)\,|f_1(s)| + a(s)\,|f_2(s)|\}\,ds$$

is uniformly convergent for large λ, *and* $\int_0^\infty |S(s,\lambda)|\,ds = O(|\lambda|^{-1})$, *then*

$$f(r) = -\lim_{R\to\infty}\frac{1}{\pi i}\int_{-R+i\epsilon}^{R+i\epsilon}\Phi(r,\lambda)\,d\lambda,$$

where $\Phi(r,\lambda) = [\Phi_1, \Phi_2]$ *is defined by*

$$\Phi_1(r,\lambda) = z_1(r,\lambda)\int_0^r \{c(s)v_1(s)f_1(s) + a(s)v_2(s)f_2(s)\}\,ds + v_1(r,\lambda)\int_r^\infty \{c(s)z_1(s)f_1(s) + a(s)z_2(s)f_2(s)\}\,ds,$$

$$\Phi_2(r,\lambda) = z_2(r,\lambda)\int_0^r \{c(s)v_1(s)f_1(s) + a(s)v_2(s)f_2(s)\}\,ds + v_2(r,\lambda)\int_r^\infty \{c(s)z_1(s)f_1(s) + a(s)z_2(s)f_2(s)\}\,ds.$$

This statement is true uniformly in ϵ.

The proof of this theorem is begun by writing Φ_1 in the form

$$\Phi_1(r,\lambda) = z_1(r,\lambda)\left[\int_0^{r-\delta} + \int_{r-\delta}^r\right] + v_1(r,\lambda)\left[\int_r^{r-\delta} + \int_{r+\delta}^\infty\right] = \Phi_{11} + \Phi_{12} + \Phi_{13} + \Phi_{14}.$$

Since $\Phi(r,\lambda)$ is holomorphic in the upper half-plane, the contour $[-R + i\epsilon, R + i\epsilon]$ may be deformed into the semicircle of radius R in the upper half-plane. First, consider the integral Φ_{14}. Because of the order properties of v_1 and z_1, this becomes

$$\Phi_{14} = v_1(r,\lambda)\int_{r+\delta}^\infty \{c(s)z_1(s)f_1(s) + a(s)z_2(s)f_2(s)\}\,ds = O\left\{k(r)e^{\mathrm{Im}\,\alpha(r)}\int_{r+\delta}^\infty e^{-\mathrm{Im}\,\alpha(s)}E(s)[c(s)\,|f_1(s)| + a(s)\,|f_2(s)|]\,ds\right\} = O\{e^{-\tau h'(r)\delta}\}.$$

Here it has been assumed that $\int_0^\infty E(s)\{c(s)\,|f_1(s)| + a(s)\,|f_2(s)|\}\,ds$ is uniformly convergent in λ for large $|\lambda|$. On the semicircle, $\lambda = Re^{i\theta}$, where $0 < \arg\lambda < \pi$ and $\tau = \mathrm{Im}\,\lambda = R\sin\theta$. Hence,

$$\left|\int \Phi_{14}\right| = \int_\epsilon^{\pi\epsilon} O\{e^{-R\sin\theta h'(r)\delta}\}R\,d\theta,$$

and this integral tends to 0 as $R \to \infty$. In a similar fashion, we can show that $\Phi_{11} \to 0$ for $R \to \infty$.

Now consider the integral

$$\Phi_{12}(r,\lambda) = z_1(r,\lambda)\int_{r-\delta}^r [c(s)v_1(s)f_1(s) + a(s)v_2(s)f_2(s)]\,ds$$

on the previous semicircle. With the use of the order properties for large λ, this becomes

$$\Phi_{12} = -\frac{1}{2}k(r)M^{-1}(\lambda) \times \exp\left[\int_r^\infty S(s)\,ds + i\alpha(r)\right][1 + O(|\lambda|^{-1})] \times \int_{r-\delta}^r \{k(r)\sin[\alpha(s) + \beta]c(s)f_1(s) + k^{-1}(r)\cos[\alpha(s) + \beta]a(s)f_2(s)\}\,ds + O\left\{\frac{e^{-\mathrm{Im}\,\alpha(r)}}{|\lambda|}\int_{r-\delta}^\delta e^{\mathrm{Im}\,\alpha(s)}[c(s)\,|f_1(s)| + a(s)\,|f_2(s)|]\,ds\right\}.$$

The last term vanishes on the semicircle as $R \to \infty$. The terms involving $O(|\lambda|^{-1})$ and $\exp i\{\alpha(s) + \beta\}$, inside the integral, also vanish on the semicircle as $R \to \infty$. The remaining term is

$$I = -\frac{1}{2}k(r)M^{-1}(\lambda)\exp\left[\int_0^\infty S(s)\,ds + i\alpha(r)\right] \times \int_{r-\delta}^\delta e^{-i[\alpha(s)+\beta]}\left\{-\frac{1}{2i}k(r)c(s)f_1(s) + \frac{1}{2}k^{-1}(r)a(s)f_2(s)\right\}ds.$$

Now it is easily shown that

$$M(\lambda) = \tfrac{1}{2}e^{-i\beta}\{1 + O(1/\lambda)\}.$$

Consequently, for large λ,

$$I = -k(r)e^{+i[\beta+\alpha(r)]}\int_{r-\delta}^r e^{-i[\alpha(s)+\beta]}\left\{-\frac{k(s)}{2i}a(s)f_1(s) + \frac{k^{-1}(s)}{2}a(s)f_2(s)\right\}ds.$$

By adding and subtracting a term and putting, for large λ,

$$\alpha(r) = \lambda h(r) + g(r),$$

this becomes

$$I \cong -k(r)e^{i[\lambda h(r)+g(r)]}\Big\{\int_{r-\delta}^{r} e^{-i[\lambda h(s)+g(r-)]} \times \Big[-\frac{k^{-1}(r-)}{2i} f_1(r-) + \frac{k(r-)}{2} f_2(r-)\Big][a(s)c(s)]^{\frac{1}{4}}\, ds + \int_{r-\delta}^{r} e^{-i\lambda h(s)}\Big(e^{-ig(s)}\Big[-\frac{k(s)}{2i} c(s)f_1(s) + \frac{k(s)}{2} a(s)f_2(s)\Big] + e^{-ig(r-)}\Big[\frac{k^{-1}(r-)}{2i} f_1(r-) + \frac{k(r-)}{2} f_2(r-)\Big]\Big)\, ds\Big\} = I_1 + I_2.$$

Consider I_1 where f_1 and f_2 are taken to be of bounded variation so that their one-sided limits exist. It is easily seen that

$$I_1 = -k(r)e^{+i\lambda h(r)}\Big[-\frac{k^{-1}(r-)}{2i} f_1(r-) + \frac{k(r-)}{2} f_2(r-)\Big] \times \int_{r-\delta}^{r} e^{-i\lambda h(s)}[a(s)c(s)]^{\frac{1}{4}}\, ds = \Big[\frac{f_1}{2i} - k^2\frac{f_2}{2}\Big]_{r-}(i\lambda^{-1})[1 - e^{+i\lambda\delta h'(r)}].$$

For fixed δ and large λ, the term $e^{ih'(r)\delta\lambda}$ when integrated over the semicircle vanishes as $R \to \infty$. Consequently,

$$\int I_1(\lambda)\, d\lambda = \frac{1}{2}\pi i[f_1(r-) - ik^2(r-)f_2(r-)].$$

Next, consider I_2 and let

$$F(s) = e^{-ig(s)}\Big[-\frac{k(s)}{2i} c(s)f_1(s) + \tfrac{1}{2}k^{-1}(s)a(s)f_2(s)\Big].$$

Hence,

$$I_2 = -k(r)e^{+i[\lambda h(r)+g(r)]} \times \int_{r-\delta}^{r} e^{-ih(s)}[a(s)c(s)]^{\frac{1}{4}}[F(s) - F(r-)]\, ds.$$

Because $f_1(r)$ and $f_2(r)$ are of bounded variation in the neighborhood of r, $F(s)$ is also. Hence,

$$|I_2| \le k(r)\int_{r-\delta}^{r} |F(s) - F(r-)| \times [a(s)c(s)]^{\frac{1}{4}}e^{-R\sin\theta[h(r)-h(s)]}\, ds = O(\delta/R\sin\theta).$$

Therefore,

$$\Big|\int_{\epsilon}^{\pi-\epsilon} I_2\, d\lambda\Big| = O(\delta) \to 0 \quad \text{for} \quad \delta \to 0.$$

The integral Φ_{12} over the semicircle as $R \to \infty$ thus contributes exactly

$$\frac{1}{2}\pi i[f_1(r-) - ik(r-)f_2(r-)].$$

Similarly, Φ_{13} contributes

$$\frac{1}{2}\pi i[f_1(r+) + ik(r+)f_2(r+)].$$

Therefore,

$$-\lim_{R\to\infty}\int_{-R+i\epsilon}^{R+i\epsilon} \Phi_1(r, \lambda)\, d\lambda = \frac{1}{2}\pi i[f_1(r-) + f_1(r+)] - \frac{1}{2}\pi[f_2(r+) - f_2(r-)]k(r).$$

Consequently, if f_1 and f_2 are continuous, it follows that uniformly in ϵ,

$$-\frac{1}{\pi i}\lim_{R\to\infty}\int_{-R+i\epsilon}^{R+i\epsilon} \Phi_1(r, \lambda)\, d\lambda = f_1(r),$$

and similarly that

$$-\frac{1}{\pi i}\lim_{R\to\infty}\int_{-R+i\epsilon}^{R+i\epsilon} \Phi_2(r, \lambda)\, d\lambda = f_2(r).$$

8. EXPANSIONS

In the previous section, the following representation for the function $f(r)$ was shown to be true:

$$f(r) = -\lim_{R\to\infty}\frac{1}{\pi i}\int_{-R+i\epsilon}^{R+i\epsilon} \Phi(r, \lambda)\, d\lambda. \tag{25}$$

The problem in this section is to investigate the behavior of the integral as $\epsilon \to 0$. It again suffices to consider only $\Phi_1(r, \lambda)$, because the same arguments apply to $\Phi_2(r, \lambda)$.

First, it is true that relation (25) may be replaced by

$$f(r) = -\lim_{R\to\infty}\frac{1}{\pi}\int_{-R+i\epsilon}^{R+i\epsilon} \operatorname{Im}\Phi(r, \lambda)\, d\lambda. \tag{26}$$

Because $\Phi(r, \lambda)$ is analytic in the lower and upper half-planes, it follows immediately from the convergence theorem that

$$f(r) = \lim_{R\to\infty}\frac{-1}{i\pi}\int_{R-i\epsilon}^{-R-i\epsilon} \Phi(r, \lambda)\, d\lambda.$$

In this expression, keep ϵ fixed and let $\lambda = \sigma - i\epsilon = \bar{\lambda} - 2i\epsilon$, so that

$$f(r) = \lim_{R\to\infty}\frac{-1}{i\pi}\int_{R+i\epsilon}^{-R+i\epsilon} \Phi(r, \bar{\lambda} - 2i\epsilon)\, d\bar{\lambda} = \lim_{R\to\infty}\frac{1}{i\pi}\int_{-R+i\epsilon}^{R+i\epsilon} \Phi(r, \bar{\lambda})\, d\bar{\lambda}.$$

If this last expression is added to the other representation for $f(r)$, we obtain

$$2f(r) = -\lim_{R\to\infty} \frac{1}{i\pi} \int_{-R+i\epsilon}^{R+i\epsilon} [\Phi(r, \lambda) - \Phi(r, \bar{\lambda})]\, d\lambda$$
$$= -\lim_{R\to\infty} \frac{2}{\pi} \int_{-R+i\epsilon}^{R+i\epsilon} \operatorname{Im}\, [\Phi(r, \lambda)]\, d\lambda,$$

because $\Phi(r, \bar{\lambda}) = \overline{\Phi(r, \lambda)}$. Formula (26) is thus verified.

Formally, because $v(r, \lambda)$ and $w(r, \lambda)$ are real for real λ, it follows that

$$\operatorname{Im} \left\{ -\frac{1}{\pi} \int_{-R+i\epsilon}^{R+i\epsilon} \Phi_1(r, \lambda)\, d\lambda \right\}$$
$$= \operatorname{Im} \left[-\frac{1}{\pi} \int_{-R+i\epsilon}^{R+i\epsilon} \left\{ [w_1(r, \lambda) + m(\lambda)v_1(r, \lambda)] \times \int_0^r [c(s)v_1(s)f_1(s) + a(s)v_2(s)f_2(s)]\, ds \right\} d\lambda \right]$$
$$+ \operatorname{Im} \left[-\frac{1}{\pi} \int_{-R+i\epsilon}^{R+i\epsilon} v_1(r, \lambda) \int_r^\infty \{ [w_1(s, \lambda) + m(\lambda)v_1(s, \lambda)]c(s)f_1(s) + [w_2(s, \lambda) + m(\lambda)v_2(s, \lambda)]a(s)f_2(s)\, ds \}\, d\lambda \right] \quad (27)$$
$$\to \frac{1}{\pi} \int_{-\infty}^{\infty} v_1(r, \lambda)\, d\rho(\lambda) \int_0^r [c(s)v_1(s)f_1(s) + a(s)v_2(s)f_2(s)]\, ds$$
$$+ \frac{1}{\pi} \int_{-\infty}^{\infty} v_1(r, \lambda) \int_r^\infty [c_1(s)v_1(s)f_1(s) + a(s)v_2(s)f_2(s)]\, ds\, d\rho(\lambda)$$
$$= \frac{1}{\pi} \int_{-\infty}^{\infty} v_1(r, \lambda)\, d\rho(\lambda) \int_0^\infty [c(s)v_1(s)f_1(s) + a(s)v_2(s)f_2(s)]\, ds.$$

In this last expression, it is assumed that $R \to \infty$, $\epsilon \to 0$, and that

$$\rho(\lambda) = -\lim_{\epsilon\to 0} \int_0^r \operatorname{Im}\, \{m(s + i\epsilon)\}\, ds.$$

In those cases which involve a continuous spectrum, it has been shown[5] that

$$\operatorname{Im}\, \{m(\lambda)\} = -1/[\mu_1^2(\sigma) + \nu_1^2(\sigma)],$$

where μ_1 and ν_1 cannot vanish simultaneously. In such cases, the contributions to the expansions become for $f_1(r)$

$$\frac{1}{\pi} \int \frac{v_1(r, \lambda)}{\mu_1^2 + \nu_1^2}\, d\lambda \int_0^\infty [c(s)v_1(s, \lambda)f_1(s) + a(s)v_2(s)f_2(s)]\, ds,$$

and for $f_2(r)$

$$\frac{1}{\pi} \int \frac{v_2(r, \lambda)}{\mu_1^2 + \nu_1^2}\, d\lambda \int_0^\infty [c(s)v_1(s, \lambda)f_1(s) + a(s)v_2(s, \lambda)f_2(s)]\, ds,$$

where the integration over λ is over the interval of the continuous spectrum.

For the λ intervals, where the only singularities of $m(\lambda)$ are poles, and a point spectrum results, the associated contributions to the expansions are series. Formally, these contributions become for $f_1(r)$,

$$\sum_n c_n v_1(r, \lambda_n),$$

and for $f_2(r)$,

$$\sum_n c_n v_2(r, \lambda_n),$$

where

$$c_n = r_n \int_0^\infty [c(s)v_1(s, \lambda_n)f_1(s) + a(s)v_2(s, \lambda_n)f_2(s)]\, ds,$$

and

$$r_n = \int_0^\infty [c(s)v_1^2(s, \lambda_n) + a(s)v_2^2(s, \lambda_n)]\, ds.$$

The functions $f_1(r)$ and $f_2(r)$, in order to be expanded over the interval from 0 to ∞, cannot be entirely independent, but must satisfy the condition

$$k^{-1}(0)f_1(0) \cos \beta + k(0)f_2(0) \sin \beta = 0,$$

if v satisfies the condition

$$k^{-1}(0)v_1(0) \cos \beta + k(0)v_2(0) \sin \beta = 0.$$

The above formalities may be made rigorous. For instance, let it be assumed that the spectrum of the system (2), (3) is continuous, and consider the right-hand side of relation (27). Because v and w are analytic functions of $\lambda = \sigma + i\tau$, and because v and w are real functions of λ for real λ, it follows that

$$\operatorname{Im}\, \{v\} = O(\epsilon) \quad \text{and} \quad \operatorname{Im}\, \{w\} = O(\epsilon),$$

as $\tau = \epsilon \to 0$. Therefore, for r, s in a fixed interval, we have

$$\operatorname{Im}\, \{z_1(r, \lambda)v_1(s, \lambda) - v_1(r, \lambda)z_1(s, \lambda)\}$$
$$= \operatorname{Im}\, \{w_1(r, \lambda)v_1(s, \lambda) - v_1(r, \lambda)w_1(s, \lambda)\}$$
$$= O(\epsilon)$$

and

$$\operatorname{Im}\, \{z_1(r)v_2(s) - v_1(r)z_2(s)\}$$
$$= \operatorname{Im}\, \{[w_1(r) + m(\lambda)v_1(r)]v_2(s) - v_1(r)[w_2(s) + m(\lambda)v_2(s)]\}$$
$$= \operatorname{Im}\, \{w_1(r)v_2(s) - v_1(r)w_2(s)\}$$
$$= O(\epsilon).$$

The right-hand side of Eq. (27) can be rewritten in the form

$$\operatorname{Im}\Big[-\frac{1}{\pi}\int_{-R+i\epsilon}^{R+i\epsilon}\Big\{\int_0^\infty v_1(r)[c(s)z_1(s)f_1(s) + a(s)z_2(s)f_2(s)]\,ds - \int_0^r v_1(r)[c(s)z_1(s)f_1(s) + a(s)z_2(s)f_2(s)]\,ds + \int_0^r z_1(r)[c(s)v_1(s)f_1(s) + a(s)v_2(s)f_2(s)]\,ds\Big\}\,d\lambda\Big].$$

With the use of the above order properties, this becomes, for fixed R and $\epsilon \to 0$,

$$\operatorname{Im}\Big[-\frac{1}{\pi}\int_{-R+i\epsilon}^{R+i\epsilon} v_2(r)\int_0^\infty [c(s)z_1(s)f_1(s) + a(s)z_2(s)f_2(s)]\,ds\,d\lambda\Big] + O(\epsilon).$$

To evaluate this expression, only the two contributing terms must be considered. The first is for fixed ϵ, given by

$$\int_{-R+i\epsilon}^{R+i\epsilon} \operatorname{Im}[v_1(r,\lambda)]\,d\lambda \int_0^\infty \operatorname{Re}[c_1(s)z_1(s)f_1(s) + a(s)z_2(s)f_2(s)]\,ds$$

$$= \int_{-R}^{R} \operatorname{Im}[v_1(r,\sigma+i\epsilon)]\,d\sigma \int_0^\infty \operatorname{Re}[c_1(s)z_1(s)f_1(s) + a(s)z_2(s)f_2(s)]\,ds$$

$$= O(\epsilon)\int_{-R}^{R} d\sigma \int_0^\infty [c(s)\,|z_1(s)f_1(s)| + a(s)\,|z_2(s)f_2(s)|]\,ds$$

$$= O(\epsilon)\Big[\int_{-R}^{R}\Big\{\int_0^\infty [c(s)\,|z_1f_1|^2 + a(s)\,|z_2f_2|^2]\,ds\Big\}\,d\sigma\Big]^{\frac{1}{2}}$$

$$= O(\epsilon)O(\epsilon^{-\frac{1}{2}}) = O(\epsilon^{\frac{1}{2}}).$$

Here Schwarz's inequality and Lemma 2 (given in Appendix) have been used. Similarly,

$$\int_{-R}^{R} \operatorname{Re}[v_1(r,\sigma+i\epsilon) - v_1(r,\sigma)]\,d\sigma \times \int_0^\infty \operatorname{Im}[c(s)z_1(s)f_1(s) + a(s)z_2(s)f_2(s)]\,ds = O(\epsilon^{\frac{1}{2}}).$$

The remaining contribution to the integral is

$$-\frac{1}{\pi}\int_{-R}^{R} v_1(r,\sigma)\,d\sigma \int_0^\infty \operatorname{Im}[c(s)z_1(s,\sigma+i\epsilon)f_1(s) + a(s)z_2(s,\sigma+i\epsilon)f_2(s)]\,ds = -\frac{1}{\pi}\int_{-R}^{R} v_1(r,\sigma)\,d\sigma \times \int_0^\infty \operatorname{Im}[c(s)f_1(s)\{z_1(s,\sigma+i\epsilon) - z_1(s,\sigma)\} + a(s)f_2(s)\{z_2(s,\sigma+i\epsilon) - z_2(s,\sigma)\}]\,ds$$

$$-\frac{1}{\pi}\int_{-R}^{R} v_1(r,\sigma)\,d\sigma \int_0^\infty \operatorname{Im}[c(s)f_1(s)z_1(s,\sigma) + a(s)f_2(s)z_2(s,\sigma)]\,ds.$$

The first integral on the right-hand side is $O(\epsilon)$. To show this, order properties are needed for $m(\lambda)$ and $z(r,\lambda)$ as $\epsilon \to 0$. For real λ, $w(r,\lambda)$ and $v(r,\lambda)$ are real, but $m(\lambda)$ may be complex. In cases of a continuous spectrum and real λ, it has been shown[5] that

$$m(\lambda) = [-(\mu_w\mu_v + \nu_w\nu_v) - i]/(\mu_v^2 + \nu_v^2),$$

where μ_v and ν_v cannot vanish simultaneously and where μ_w, μ_v, ν_w, ν_v are real. For complex $\lambda = \sigma + i\tau$, the corresponding μ_w, μ_v, ν_w, ν_v tend to real values for real λ in the manner of $\mu_w(\sigma + i\epsilon) = \mu_w(\sigma) + O(\epsilon)$ for $\epsilon = \tau \to 0$. Consequently, for $\epsilon = \tau \to 0$,

$$m(\sigma + i\epsilon) = -\frac{[\mu_w(\sigma)\mu_v(\sigma) + \nu_w(\sigma)\nu_v(\sigma)] + i}{\mu_v^2(\sigma) + \nu_v^2(\sigma)}\,[1 + O(\epsilon)].$$

It now follows directly that

$$\operatorname{Im}[z_1(s,\sigma+i\epsilon) - z_1(s,\sigma)] = \operatorname{Im}[\{w_1(s,\sigma+i\epsilon) - w_1(s,\sigma)\} + m(\sigma+i\epsilon) \times \{v_1(s,\sigma+i\epsilon) - v_1(s,\sigma)\} + v_1(s,\sigma)\{m(\sigma+i\epsilon) - m(\sigma)\}] = O(\epsilon).$$

Consequently, the only term that remains for fixed R and $\epsilon \to 0$ is

$$-\frac{1}{\pi}\int_{-R}^{R} v_1(r,\sigma)\,d\sigma \int_0^\infty \operatorname{Im}[c(s)f_1(s)z_1(s,\sigma) + a(s)f_2(s)z_2(s,\sigma)]\,ds.$$

Hence, for a given R and $\epsilon \to 0$, this becomes

$$\frac{1}{\pi}\int_{-R}^{R} v_1(r,\sigma)\,d\sigma\,[\mu_v^2(\sigma) + \nu_v^2(\sigma)]^{-1} \times \int_0^\infty [c(s)f_1(s)v_1(s,\sigma) + a(s)f_2(s)v_2(s,\sigma)]\,ds.$$

Thus, using the convergence theorem and this result, the expansion for $f_1(r)$ is given by

$$f_1(r) = \lim_{R\to\infty}\Big[-\frac{1}{\pi}\int_{-R+i\epsilon}^{R+i\epsilon} \operatorname{Im}\,\Phi_1(r,\lambda)\,d\lambda\Big] = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{v_1(r,\sigma)\,d\sigma}{\mu_v^2(\sigma) + \nu_v^2(\sigma)} \times \int_0^\infty [c(s)f_1(s)v_1(s,\sigma) + a(s)f_2(s)v_2(s,\sigma)]\,ds.$$

The expansion for $f_2(r)$ can be proved similarly.

APPENDIX

Lemma 1. *For any fixed* λ *and* λ', $\lim_{r\to\infty} W[z(r, \lambda), z(r, \lambda')] = 0$.

Proof: Because $z(r, \lambda)$ satisfies at r_1 a boundary condition which is independent of λ,

$$W_{r_1}[w(r, \lambda) + l(\lambda)v(r, \lambda), w(r, \lambda') + l(\lambda)v(r, \lambda')] = 0,$$

$$W_{r_1}[z(r, \lambda) + [l(\lambda) - m(\lambda)]v(r, \lambda), z(r, \lambda') + [l(\lambda') - m(\lambda')v(r, \lambda')] = 0,$$

$$\begin{aligned} W_{r_1}[z(r, \lambda), z(r, \lambda')] &+ [l(\lambda) - m(\lambda)]W_{r_1}[v(r, \lambda), z(r, \lambda')] \\ &+ [l(\lambda') - m(\lambda')]W_{r_1}[z(r, \lambda), v(r, \lambda')] \\ &+ [l(\lambda) - m(\lambda)][l(\lambda') - m(\lambda')] \\ &\times W_{r_1}[v(r, \lambda), v(r, \lambda')] = 0. \end{aligned}$$

Now,

$$W_{r_1}[v(r, \lambda), z(r, \lambda')] = (\lambda - \lambda') \times \int_0^{r_1} [a(s)z_2(s)v_2(s) + c(s)z_1(s)v_1(s)]\, ds$$

$$+ W_0[v, z] = O\left[\int_0^{r_1} |a(s)z_2(s)v_2(s) + c(s)v_1(s)z_1(s)|\, ds\right] + o(1) = O\left[\int_0^{r_1} \{a(s)\,|z_2(s)|^2 + c(s)\,|z_1(s)|^2\} \int_0^{r_1} \{a(s)\,|v_2(s)|^2 + c(s)\,|v_1(s)|^2\}\right]^{\frac{1}{2}}$$

$$+ o(1)O\left[\int_0^{\infty} \{a(s)\,|v_2(s)|^2 + c(s)\,|v_1(s)|^2\}^{\frac{1}{2}}\right] + o(1),$$

as $r \to \infty$. Here, one has used the fact that z is $L^2(0, \infty)$.

In the limit-point case,

$$|l(\lambda) - m(\lambda)| \leq 2R_{r_1} = \left[\int_0^b \{c(s)\,|v_1(s)|^2 + a(s)\,|v_2(s)|^2\}\right]^{-1}.$$

Consequently,

$$\lim_{r_1\to\infty} |l(\lambda) - m(\lambda)|\, W_{r_1}[v(r, \lambda), z(r, \lambda)] = 0.$$

Similar arguments apply for the other terms involving $[l(\lambda) - m(\lambda)]$. In the limit-circle case, $\int_0^{r_1} [c(s)\,|v_1(s)|^2 + a(s)\,|v_2(s)|^2]\, ds$ is bounded but $l(\lambda) \to m(\lambda)$. The lemma therefore follows.

Lemma 2.

$$\int_0^{\infty} [c(s)\,|z_1(s, \lambda)|^2 + a(s)\,|z_2(s, \lambda)|^2]\, ds = -\frac{\operatorname{Im}[m(\lambda)]}{\tau}.$$

Proof:

$$(\lambda' - \lambda) \int_0^{r_1} [a(s)z_2(s, \lambda)z_2(s, \lambda') + c(s)z_1(s, \lambda)z_1(s, \lambda')]\, ds = W_{r_1}[z(r, \lambda), z(r, \lambda')] - W_0[z(r, \lambda), z(r, \lambda')].$$

From Lemma 1, the term $W_{r_1} \to 0$ as $b \to \infty$. Also, since $z = w + mv$,

$$W_0[z(r, \lambda), z(r, \lambda')] = z_1(r, \lambda)z_2(r, \lambda') - z_2(r, \lambda)z_1(r, \lambda') = m(\lambda) - m(\lambda'),$$

and the lemma follows immediately.

Kinetic Equation for an Unstable Plasma

R. BALESCU
Faculté des Sciences, Université Libre de Bruxelles, Brussels, Belgium
(Received 19 April 1963)

A kinetic equation is derived for the description of the evolution in time of the distribution of velocities in a spatially homogeneous ionized gas which, at the initial time, is able to sustain exponentially growing oscillations. This equation is expressed in terms of a functional of the distribution function which obeys the same integral equation as in the stable case. Although the method of solution used in the stable case breaks down, the equation can still be solved in closed form under unstable conditions, and hence an explicit form of the kinetic equation is obtained. The latter contains the "normal" collision term and a new additional term describing the stabilization of the plasma. The latter acts through friction and diffusion and brings the plasma into a state of neutral stability. From there on the system evolves towards thermal equilibrium under the action of the normal collision term as well as of an additional Fokker–Planck-like term with time-dependent coefficients, which however becomes less and less efficient as the plasma approaches equilibrium.

1. INTRODUCTION

In recent years many convergent efforts have been directed towards the study of the evolution of ionized gases under various conditions. The role of the long-range Coulomb interactions in driving such systems towards thermal equilibrium is one of the most interesting aspects of this problem. In 1960 the author[1,2] and, independently, Lenard[3] and Guernsey,[4] derived a kinetic equation describing the evolution of a spatially homogeneous electron gas towards equilibrium:

$$\partial_t \varphi(\mathbf{v};\, t) = \mathcal{C}\{\varphi\} \tag{1.1}$$

$$\mathcal{C}\{\varphi\} = 2e^4 cm^{-2} \int d\mathbf{k} \int d\mathbf{v}_1\, \mathbf{k}\cdot\boldsymbol{\partial} \frac{\delta(\mathbf{k}\cdot\mathbf{v} - \mathbf{k}\cdot\mathbf{v}_1)}{k^4 \,|\, \epsilon^-(\mathbf{k}\cdot\mathbf{v}/k)|} \times \mathbf{k}\cdot(\boldsymbol{\partial} - \boldsymbol{\partial}_1)\varphi(\mathbf{v})\varphi(\mathbf{v}_1), \tag{1.2}$$

with the following meanings for the symbols: e is the charge of the electron, c the number density, m the mass, $\varphi(\mathbf{v};\, t)$ the velocity distribution function for one particle,

$$\boldsymbol{\partial} \equiv \partial/\partial\mathbf{v}; \qquad \boldsymbol{\partial}_i \equiv \partial/\partial\mathbf{v}_i,$$

$$\epsilon^-(\nu) = 1 + (\omega_p^2/k^2)\pi i \int d\mathbf{v}_1\, \delta_-(l\nu - \mathbf{l}\cdot\mathbf{v}_1)\mathbf{l}\cdot\boldsymbol{\partial}_1\varphi(\mathbf{v}_1),$$

$$\omega_p^2 = 4\pi e^2 c/m.$$

The singular functions $\delta_\pm(x)$ are defined as follows:

$$\delta_\pm(x) = \delta(x) \pm i\pi^{-1}\mathcal{P}(1/x), \tag{1.3}$$

where $\mathcal{P}$ denotes the Cauchy principal part. The highly nonlinear equation (1.1) [$\epsilon^-(\nu)$ is a functional of $\varphi(\mathbf{v})$!] explicitly takes into account the collective nature of the collision processes. The ideas leading to Eq. (1.1) have been exploited and extended in various directions (see for instance references 2; 5–8). The theory has however an important limitation which will be presently discussed.

It is well known from the theory of plasma oscillations [2,9–12] that a spatial disturbance at the initial time can behave in two distinct ways at later times: in so-called "stable" plasmas it will give rise to damped oscillations, whereas in "unstable" plasmas there appear exponentially amplified oscillations. This behavior results from the properties of the linearized Vlassov equation of the plasma. The occurrence of one or the other case depends only on the initial velocity distribution function $\varphi(\mathbf{v})$ and not on the initial perturbation. More precisely, in the linearized Vlassov approximation, the intrinsic behavior of the plasma is uniquely characterized by a single function, called the *dielectric constant* $\epsilon_{\mathbf{k}}^+(w)$, which is defined for complex values of the frequency kw by

$$\epsilon_{\mathbf{k}}^+(w) = 1 - \frac{\omega_p^2}{k^2} \int d\mathbf{v} \frac{\mathbf{k}\cdot\boldsymbol{\partial}\varphi(\mathbf{v})}{\mathbf{k}\cdot\mathbf{v} - kw}, \quad w \in S_+. \tag{1.4}$$

This definition, valid when w lies in the upper half-plane S_+, is extended to the whole complex plane by analytic continuation. The proper frequencies of the plasma are the roots w_i of the equation

[1] R. Balescu, Phys. Fluids **3**, 52 (1960).
[2] R. Balescu, *Statistical Mechanics of Charged Particles* (Interscience Publishers, Inc., New York, to be published).
[3] A. Lenard, Ann. Phys., **3**, 390 (1960).
[4] R. L. Guernsey, *The Kinetic Theory of Fully Ionized Gases*, U. S. Office of Naval Research, Contract No. Nonr. 1224(15), July 1960.
[5] R. Balescu and H. S. Taylor, Phys. Fluids **4**, 85 (1961).
[6] R. Balescu, Phys. Fluids **4**, 94 (1961).
[7] R. L. Guernsey, Phys. Fluids **5**, 322 (1962).
[8] N. Rostoker and M. Rosenbluth, Phys. Fluids **3**, 1 (1960).
[9] L. Landau, J. Phys. (USSR) **10**, 25 (1946).
[10] M. E. Ghertsenstein, J. Exptl. Theoret. Phys. (USSR) **23**, 669 (1952).
[11] J. D. Jackson, J. Nucl. Energy **C1**, 171 (1960).
[12] O. Penrose, Phys. Fluids **3**, 258 (1960).

$\epsilon_{\mathbf{k}}^{+}(w_i) = 0$. If there is a root lying in the upper half-plane, the plasma is unstable.

Coming back to Eq. (1.1), it has been shown that a necessary condition for its validity is that the distribution $\varphi(\mathbf{v})$ describes initially (and hence at all later times) a stable plasma. Its derivation breaks down for unstable plasmas.[2,5,13]

It is physically obvious that the exponential growth of plasma oscillations cannot be unlimited. The mechanism which stops this growth has been assumed by many authors to be found in the nonlinear Vlassov equation. However in a recent paper, Drummond and Pines[14] have shown that, besides the latter mechanism, there exists another one within the framework of the linear Vlassov theory. These authors demonstrate that the very existence of unstable oscillations induces a change in time of the momentum distribution; this change is such as to finally suppress the instability.

However, this mechanism depends crucially on the existence of periodic inhomogeneities in the plasma. In a spatially homogeneous system or in a state with localized inhomogeneities this mechanism does not exist. Our purpose in the present paper is to show that the instability induces also another type of evolution of the velocity distribution. This new mechanism is effective in all situations, because it is determined by the behavior of the binary correlations. In the present paper we will however limit ourselves to the study of spatially homogeneous plasmas. The crucial point in the derivation of our equation is the following. It will be shown that the two-body correlation function undergoes in an unstable homogeneous plasma an exponential growth which has the same origin as the growth of plasma oscillations. But within the ring approximation, the rate of change of the momentum distribution is related to the binary correlations; hence the instability will induce a new term into the kinetic equation of the plasma.

Section 2 is devoted to a review of the general theory of nonequilibrium statistical mechanics and to the derivation of a formal kinetic equation. The latter is expressed in terms of a fundamental functional of the distribution $F_{\mathbf{k}}(\mathbf{v}; w)$, which obeys a certain integral equation. Although the same equation appears in the theory of stable plasmas, its mathematical properties change significantly when the plasma is unstable. After a short review of the stability conditions in Sec. 3, the stable solution is briefly discussed in Sec. 4. Section 5 is devoted to the solution of the integral equation in the unstable case. The kinetic equation is derived in Sec. 6 and some of its properties are discussed in the final section.

2. DERIVATION OF FORMAL EQUATIONS

The system studied in the present paper is the simplest idealized type of plasma which could exhibit the phenomena we have in view. It consists of a spatially homogeneous electron gas of density c imbedded in a continuous neutralizing positive background. The relevant reduced distribution functions which characterize that system are the one-particle velocity distribution $\varphi(\mathbf{v}_\alpha; t)$ and the two-body distribution, which is Fourier-analyzed as follows:

$$f_2(\mathbf{x}_\alpha - \mathbf{x}_\beta, \mathbf{v}_\alpha, \mathbf{v}_\beta; t) = c^2\Big\{\varphi(\mathbf{v}_\alpha; t)\varphi(\mathbf{v}_\beta; t) + \int d\mathbf{k}\, e^{i\mathbf{k}\cdot(\mathbf{x}_\alpha - \mathbf{x}_\beta)t}\rho_{\mathbf{k},-\mathbf{k}}(\mathbf{v}_\alpha, \mathbf{v}_\beta; t)\Big\}. \tag{2.1}$$

These functions have been defined in references 2, 15, and 16 as integrals of the fundamental N-body distribution function $f_N(\mathbf{x}_1, \cdots, \mathbf{x}_N, \mathbf{v}_1, \cdots, \mathbf{v}_N; t)$ which obeys the Liouville equation (we consider here only classical systems).

We shall now derive a kinetic equation for the one-particle function $\varphi(\mathbf{v}_\alpha; t)$ starting from the Liouville equation. The calculations of this section do not depend on the stability criterion of the plasma. We use the theory developed by Prigogine and the author,[15] generalized by Prigogine and Résibois[16,17] and by the author.[2] We sketch here the main lines of the general theory for the facility of the exposition; more details can be found in reference 2, and, with somewhat different notations, in references 16, and 17. The Liouville equation is written as follows:

$$\mathfrak{L}f_N \equiv (\mathfrak{L}_0 + e^2\mathfrak{L}')f_N = 0, \tag{2.2}$$

with

$$\mathfrak{L}_0 = \partial_t + \sum_i \mathbf{v}_i\cdot(\partial/\partial\mathbf{x}_i), \tag{2.3}$$

$$\mathfrak{L}' \equiv \sum_{i<n}\mathfrak{L}'_{in} = -\sum_{i<n} m^{-1}(\partial V_{in}/\partial\mathbf{x}_i)\cdot(\partial_i - \partial_n), \tag{2.4}$$

and

$$V_{in} = (|\mathbf{x}_i - \mathbf{x}_n|)^{-1} = \frac{1}{2\pi^2}\frac{8\pi^3}{\Omega}\sum_{\mathbf{k}}\frac{e^{i\mathbf{k}\cdot(\mathbf{x}_i - \mathbf{x}_n)}}{k^2}. \tag{2.5}$$

[13] A. Lenard, Bull. Am. Phys. Soc. 6, 189 (1961).
[14] W. E. Drummond and D. Pines, Proc. Conf. Plasma Phys. Controlled Nucl. Fusion Resch., Salzburg, 1961 (to be published).
[15] I. Prigogine and R. Balescu, Physica **25**, 281, 302 (1959).
[16] I. Prigogine, *Non-Equilibrium Statistical Mechanics*, (Interscience Publishers, Inc., New York, 1963).
[17] I. Prigogine and P. Résibois, Physica **27**, 629 (1961).

Let $\mathcal{G}(xvt \mid x'v't')$ be the retarded Green's function of the Liouville equation (2.2) [x, v stand for the set $\mathbf{x}_1, \cdots, \mathbf{x}_N, \mathbf{v}_1, \cdots, \mathbf{v}_N$; $\delta(x - x')$ means $\prod_i \delta(\mathbf{x}_i - \mathbf{x}_i')$.] This function is defined as the solution of the equation

$$\mathcal{L}\mathcal{G}(xvt \mid x'v't') = \delta(x - x')\delta(v - v')\delta(t - t'), \quad (2.6)$$

with the causal condition

$$\mathcal{G}(xvt \mid x'v't') = 0 \quad \text{for} \quad t < t'. \quad (2.7)$$

In terms of the Green's function, the solution of the initial-value problem for the Liouville equation is

$$f_N(x, v; t) = \int dx'\, dv'\, \mathcal{G}(xvt \mid x'v'0) f_N(x', v'; 0)$$
$$\equiv \mathcal{G}(t \mid 0) f_N(0). \quad (2.8)$$

The last equation defines $\mathcal{G}(t \mid t')$ as an *operator* acting on $f_N(t')$. As the Liouville operator has time-independent coefficients, it can be shown that $\mathcal{G}(t \mid t')$ is a function only of $\tau = t - t'$; moreover it vanishes for $\tau < 0$. It is therefore natural to introduce the Laplace transform of the Green's operator, called the *resolvent operator*:

$$\mathcal{R}(z) = \int_0^\infty d\tau\, e^{iz\tau} \mathcal{G}(\tau). \quad (2.9)$$

Expressing $\mathcal{G}(\tau)$ in terms of $\mathcal{R}(z)$ and substituting into (2.8), we obtain

$$f_N(t) = (2\pi)^{-1} \int_c dz\, e^{-izt} \mathcal{R}(z) f_N(0). \quad (2.10)$$

The contour C is a straight line parallel to the real axis and lying above all singularities of $\mathcal{R}(z)$.

We now introduce perturbation theory. Calling $\mathcal{R}_0(z)$ the resolvent of the unperturbed Liouville operator $\mathcal{L}_0$ [Eq. (2.3)], it can be shown that $\mathcal{R}(z)$ obeys the following equation:

$$\mathcal{R}(z) = \mathcal{R}_0(z) - e^2 \mathcal{R}_0(z) \mathcal{L}' \mathcal{R}(z). \quad (2.11)$$

The latter can be solved by successive iterations and the result, substituted in (2.10) yields

$$f_N(t) = (2\pi)^{-1} \sum_{n=0}^{\infty} \int_c dz\, e^{-izt} (-e^2)^n$$
$$\times \mathcal{R}_0(z) [\mathcal{L}' \mathcal{R}_0(z)]^n f_N(0). \quad (2.12)$$

We now go over to the Fourier representation of references 2, 15, and 16, whereby this equation is transformed into

$$\rho_{\{k\}}(v; t) = (2\pi)^{-1} \sum_{n=0}^{\infty} \int_c dz\, e^{-izt} (-e^2)^n \sum_{\{k'\}} (8\pi^3/\Omega)^{\nu'-\nu}$$
$$\times \langle \{k\} \mid \mathcal{R}_0(z) [\mathcal{L}' \mathcal{R}_0(z)]^n \mid \{k'\} \rangle \rho_{\{k'\}}(v; 0), \quad (2.13)$$

where ν and ν' are the numbers of independent nonvanishing wave vectors in the sets $\{k\}$ and $\{k'\}$. It is moreover easily shown by a direct calculation that

$$\langle \{k\} \mid \mathcal{R}_0(z) \mid \{k'\} \rangle = \frac{1}{i(\sum \mathbf{k}_n \cdot \mathbf{v}_n - z)} \prod_i \delta(\mathbf{k}_i - \mathbf{k}_i'). \quad (2.14)$$

The terms of the perturbation series (2.13) are represented by exactly the same diagrams as in reference 15; the only difference in the interpretation is that the lines in the diagrams no longer represent oscillating exponentials, but factors $\mathcal{R}_0(z)$ (which are the Laplace transforms of the former).

After this general outline, we come back to the specific problem we have in mind. We want a description of the plasma in the *"ring approximation"* introduced in references 1 and 2. It has been shown there that, in this approximation, the contributions to $\rho_0(t)$ coming from $\rho_{\{k'\}}(0)$ where $\{k'\} \neq \{0\}$, are negligible for sufficiently long times ($t \gg \omega_p^{-1}$). Hence, Eq. (2.13) for $\{k\} = \{0\}$ becomes

$$\rho_0(v; t) = (2\pi)^{-1} \sum_{n=0}^{\infty} \int_c dz\, e^{-izt} (-e^2)^n$$
$$\times \langle 0 \mid \mathcal{R}_0(z) [\mathcal{L}' \mathcal{R}_0(z)]^n \mid 0 \rangle \rho_0(v; 0). \quad (2.15)$$

It results from the structure of the perturbation series that the most general contribution to the 0–0 matrix element of the resolvent is represented by a succession of diagonal fragments, i.e. transitions from zero to zero through states in all of which there are at least two nonvanishing wave vectors. Within the ring approximation, we are interested only in diagonal fragments of a special type, called *rings*, a typical one being represented in Fig. 1,

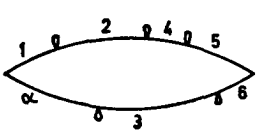

FIG. 1. A typical ring diagram.

and we reject from (2.15) the contributions of all other types of diagonal fragments. Let us introduce the notation

$$R(z) = \sum_{\substack{n=1 \\ \text{(all rings)}}}^{\infty} (-e^2)^{n+1} \langle 0 \mid \mathcal{L}' [\mathcal{R}_0(z) \mathcal{L}']^n \mid 0 \rangle. \quad (2.16)$$

The summation is carried out over all ring diagrams, i.e. all the diagrams summed in reference 1. Then the series in Eq. (2.15) can be rewritten as follows:

$$\rho_0(v; t) = (2\pi)^{-1} \sum_{m=0}^{\infty} \int_c dz\, e^{-izt} \frac{1}{-iz} \left\{ R(z) \frac{1}{-iz} \right\}^m \rho_0(v; 0)$$
$$\equiv (2\pi)^{-1} \int_c dz\, e^{-izt} \tilde{\rho}_0(v; z). \quad (2.17)$$

By time differentiation of Eq. (2.17) and integration over all velocities but $\mathbf{v}_\alpha$, we easily obtain the kinetic equation for the reduced velocity distribution function $\varphi(\mathbf{v}_\alpha; t)$,

$$\partial_t \varphi(\mathbf{v}_\alpha; t) = (2\pi)^{-1} \int_{(\alpha)} (d\mathbf{v})^{N-1} \times \int_c dz\, e^{-izt} R(z) \tilde{\rho}_0(v; z). \quad (2.18)$$

We now substitute for $\tilde{\rho}_0(v; z)$ its Laplace transform in order to come back to the time-dependent velocity distribution:

$$\partial_t \varphi(\mathbf{v}_\alpha; t) = (2\pi)^{-1} \int_c dz\, e^{-izt} \times \int_{(\alpha)} (d\mathbf{v})^{N-1} R(z) \int_0^\infty d\tau\, e^{iz\tau} \rho_0(v; \tau),$$

or, making use of the causality condition (2.7),

$$\partial_t \varphi(\mathbf{v}_\alpha; t) = (2\pi)^{-1} \int_0^t d\tau \times \int_c dz\, e^{-iz\tau} \int_{(\alpha)} (d\mathbf{v})^{N-1} R(z) \rho_0(v; t - \tau). \quad (2.19)$$

The transformation leading from (2.18) to (2.19) is a very important step for the following reason: A basic assumption in nonequilibrium statistical mechanics is the factorization of all s-particle velocity distribution functions at some initial time, a property which can then be shown to persist at all later times;

$$\varphi_s(\mathbf{v}_1, \cdots, \mathbf{v}_s; t) = \prod_{i=1}^{s} \varphi(\mathbf{v}_i; t). \quad (2.20)$$

This property implies however that the Laplace transform $\tilde{\varphi}_s(\mathbf{v}_1, \cdots, \mathbf{v}_s; z)$ is *not factorizable* [a Laplace transformation transforms (2.20) into a convolution]. But the factorization theorem is crucial in the summation procedure of the rings. By making the Laplace transformation leading to (2.19), we recover time-dependent velocity distributions, which can be factorized.

Equation (2.19) is a typical *non-Markoffian equation*: the evolution of the distribution function at time t depends on the whole past history of the system, from time zero up to the present time t. This non-Markoffian character (i.e. the memory of the past) is typical of any short-time kinetic equation.[16,17]

In the limiting case, where the duration of the memory is very short compared with the rate of change of the distribution function, (2.19) can be approximated as follows: The variation of $\rho_0(v; t - \tau)$ during the effective duration of the memory is neglected and this function is replaced by $\rho_0(v; t)$. Hence, Eq. (2.19) reduces to the Markoffian equation

$$\partial_t \varphi(\mathbf{v}_\alpha; t) = \frac{1}{2\pi} \int_c dz \frac{e^{-izt}}{z} \times \int_{(\alpha)} (d\mathbf{v})^{N-1} R(z) \rho_0(v; t). \quad (2.21)$$

If, moreover, the plasma is not close to instability, the only relevant contribution to the integral is the residue in $z = 0+$, and the right-hand side (r. h. s.) reduces to

$$\int_{(\alpha)} (d\mathbf{v})^{N-1} R(0+) \rho_0(v; t). \quad (2.22)$$

In this quantity, the summation over the rings expressed by Eq. (2.16) has been performed explicitly in references 1 and 2; it eventually leads to Eq. (1.1). We note that the summation procedure of reference 1, Sec. 3 (or of reference 2, Sec. 40) depends only on the topological structure of the rings, and can be taken over with only trivial changes for the evaluation of the more general quantity

$$\int_{(\alpha)} (d\mathbf{v})^{N-1} R(z) \rho_0(v; t - \tau)$$

appearing in the non-Markoffian Eq. (2.22). One merely has to make the substitutions

$$\varphi(\mathbf{v}; t) \to \varphi(\mathbf{v}; t - \tau),$$

$$\pi \delta_-(\mathbf{k}\cdot\mathbf{v}_j - \mathbf{k}\cdot\mathbf{v}_n) \to \frac{1}{i(\mathbf{k}\cdot\mathbf{v}_j - \mathbf{k}\cdot\mathbf{v}_n - z)}$$

in every expression occurring in the former theory. We therefore immediately quote the result of the summation over the rings:

$$\partial_t \varphi(\mathbf{v}; t) = -\left(\frac{\omega_p^2}{2\pi}\right) \times \int_c dz \int_0^t d\tau\, e^{-iz\tau} \int d\mathbf{k}\, k^{-2} i\mathbf{k}\cdot\partial F_{\mathbf{k}}(\mathbf{v}; z/k).$$

Of course, $F_{\mathbf{k}}(\mathbf{v}; z/k)$ is also a function of τ, but for the sake of simplicity, we do not write down this dependence explicitly. Permuting the z and k integrations for later convenience, and writing $w = z/k$, we obtain

$$\partial_t \varphi(\mathbf{v}; t) = -\omega_p^2 \int d\mathbf{k}\, i\mathbf{k}\cdot\partial\, \mathfrak{F}_{\mathbf{k}}(\mathbf{v}; t), \quad (2.23)$$

with

$$\mathfrak{F}_{\mathbf{k}}(\mathbf{v}; t) = \frac{1}{2\pi} \int_c dw \int_0^t d\tau\, e^{-ikw\tau} F_{\mathbf{k}}(\mathbf{v}; w). \quad (2.24)$$

The fundamental function $F_{\mathbf{k}}(\mathbf{v}; w)$ is the solution of the following integral equation:

$$\epsilon_{\mathbf{k}}^{-}(\nu - w)F_{\mathbf{k}}(\mathbf{v}; w)$$

$$= d_{\mathbf{k}}(\mathbf{v}) \int d\mathbf{v}_1 \frac{F_{-\mathbf{k}}(\mathbf{v}_1; w)}{\mathbf{k}\cdot\mathbf{v} - kw - \mathbf{k}\cdot\mathbf{v}_1} + q_{\mathbf{k}}(\mathbf{v}; w), \quad (2.25)$$

where

$$\epsilon_{\mathbf{k}}^{-}(\nu) = 1 + \frac{\omega_p^2}{k^2} \int d\mathbf{v}_1 \frac{\mathbf{k}\cdot\partial_1\varphi(\mathbf{v}_1; t - \tau)}{k\nu - \mathbf{k}\cdot\mathbf{v}_1}, \quad (2.26)$$

$$d_{\mathbf{k}}(\mathbf{v}) = (\omega_p^2/k^2)\mathbf{k}\cdot\partial\varphi(\mathbf{v}; t - \tau), \quad (2.27)$$

$$q_{\mathbf{k}}(\mathbf{v}; w) = \frac{\omega_p^2}{8\pi^3 ck^2}$$

$$\times \int d\mathbf{v}_1 \frac{\mathbf{k}\cdot(\partial - \partial_1)\varphi(\mathbf{v}; t - \tau)\varphi(\mathbf{v}_1; t - \tau)}{\mathbf{k}\cdot\mathbf{v} - kw - \mathbf{k}\cdot\mathbf{v}_1}. \quad (2.28)$$

Before solving this equation we shall briefly review some properties of the dielectric constant of the plasma.

3. CRITERION OF STABILITY

The stability properties of the plasma are completely determined by the analytical properties of the dielectric constant $\epsilon_{\mathbf{k}}^{+}(w)$ defined by Eq. (1.4). We can rewrite this equation in a simpler form by making use of the "barred" functions introduced in reference 1. The barring operation associates with every function of the vector $\mathbf{v}$ and with every wave vector $\mathbf{k}$ a barred function $\bar{f}(\nu)$, which is the integral of the original function $f(\mathbf{v})$ over the components of $\mathbf{v}$ perpendicular to $\mathbf{k}$:

$$\bar{f}(\nu) \equiv \int d\mathbf{v}\ \delta(\nu - \mathbf{k}\cdot\mathbf{v}/k)f(\mathbf{v}). \quad (3.1)$$

Writing also

$$\bar{f}'(\nu) \equiv \partial\bar{f}(\nu)/\partial\nu, \quad (3.2)$$

we can re-express the dielectric constant as follows (dropping the subscript $\mathbf{k}$):

$$\epsilon^{+}(w) = 1 - \frac{\omega_p^2}{k^2} \int_{-\infty}^{\infty} d\nu \frac{\bar{\varphi}'(\nu)}{\nu - w}, \quad w \in S_+. \quad (3.3)$$

The analytical properties of this function result from those of the Cauchy integral

$$\Phi(w) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} d\nu \frac{f(\nu)}{\nu - w}. \quad (3.4)$$

It is well known that such a function is regular for w lying in the upper half-plane S_+, and tends to a definite limit $\Phi^{(+)}(w)$ for w approaching the real axis from above. This limiting value is given by

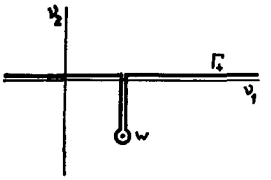

FIG. 2. Contour of integration Γ_+ for the analytic continuation of $\Phi^{(+)}(w)$ into S_-.

$$\Phi^{(+)}(w) = \frac{1}{2} \int_{-\infty}^{\infty} d\nu\ \delta_+(w - \nu)f(\nu), \quad w \text{ real}, \quad (3.5)$$

where $\delta_+(x)$ has been defined in (1.3). Moreover formula (3.4) defines, for $w \in S_-$, a function regular in S_- which, for w approaching the real axis from below, tends towards

$$\Phi^{(-)}(w) = -\frac{1}{2} \int_{-\infty}^{\infty} d\nu\ \delta_-(w - \nu)f(\nu), \quad w \text{ real}. \quad (3.6)$$

Thus, the Cauchy integral (3.4) has a discontinuity along the real axis, the jump being given by the Plemelj formula

$$\Phi^{(+)}(w) - \Phi^{(-)}(w) = f(w). \quad (3.7)$$

Thus, $\Phi(w)$ can be regarded as a two-valued analytical function, for which the real axis is a cut. The branch $\Phi^{(+)}(w)$ is defined by Eq. (3.4) for $w \in S_+$ and by Eq. (3.5) for w real. It can, in general, be continued analytically into S_- by taking the integral (3.4) on a different contour Γ_+ shown in Fig. 2, instead of the real axis

$$\Phi^{(+)}(w) = \frac{1}{2\pi i} \int_{\Gamma_+} d\nu \frac{f(\nu)}{\nu - w}, \quad w \in S_-. \quad (3.8)$$

Similar considerations obviously hold for the branch $\Phi^{(-)}(w)$.

After this brief mathematical review, we conclude that the dielectric constant $\epsilon^{+}(w)$ is a regular function of w in the upper half-plane, but will have singularities in the lower half-plane.

It is well-known in the theory of plasma oscillations that if $\epsilon^{+}(w)$ has zeros in the upper half-plane, the plasma is unstable. In the present paper we will consider only unstable plasmas of the following simple type: We assume that the velocity distribution is such that the dielectric constant has a single zero ζ_+ in the upper half-plane

$$\zeta_+ = w_0 + i\gamma_0, \quad \gamma_0 > 0 \quad \text{(unstable)}. \quad (3.9)$$

We will therefore write

$$\epsilon^{+}(w) = (w - \zeta_+)\sigma^{+}(w), \quad (3.10)$$

where $\sigma^{+}(w)$ is regular and different from zero in S_+.

We will also encounter a "minus-dielectric constant" $\epsilon^{-}(w)$; it is precisely this function which

enters Eq. (2.25) as the coefficient of $F_{\mathbf{k}}(\mathbf{v}; w)$. $\epsilon^-(w)$ is defined by formula (3.3) for $w \in S_-$. It is therefore regular in the lower half-plane and has singularities in S_+. It satisfies the relation

$$\epsilon^-(w) = [\epsilon^+(w^*)]^*, \tag{3.11}$$

the star denoting complex conjugation. Hence for an unstable plasma the minus dielectric constant has a zero in S_-;

$$\epsilon^-(w) = (w - \zeta_-)\sigma^-(w), \tag{3.12}$$

with

$$\zeta_- = \zeta_+^* = w_0 - i\gamma_0. \tag{3.13}$$

Before concluding this section, the following remark is in order: The function $\epsilon^+(w)$ has a (possibly infinite) set of zeros ζ_i in the lower half-plane. In normal situations the closest of these zeros are, roughly speaking, at an average distance $\omega_p/\bar{k}$ from the real axis ($\bar{k}$ being some average wave vector). This distance will be considered as large, because ω_p^{-1} measures the short time scale in the ring approximation (see reference 2). As the plasma is stabilized the zero, ζ_+ moves down towards the real axis and eventually crosses it; from here on the plasma is stable. However, the zero ζ_+ still retains a distinct property as compared with the other zeros ζ_i (which are presumably not much affected by the stabilization process): it remains close to the real axis for a certain time. This remark will prove important in Sec. 6.

4. SOLUTION OF THE INTEGRAL EQUATION (2.25) IN THE STABLE CASE

In order to evaluate the integral in Eq. (2.24), we must know the function $F_{\mathbf{k}}(\mathbf{v}; w)$ for values of w lying on the contour C of the inverse Laplace transformation, i.e. above all the singularities of $F_{\mathbf{k}}$. Physically, we expect that in the stable case $F_{\mathbf{k}}(\mathbf{v}; w)$ has no singularities (in w) located in the upper half-plane. If this is true, we can pull down the contour C and bring it on the upper edge of the real axis. A detailed justification of this process will be found in the next section. The advantage of this operation is that the Fredholm equation (2.25) becomes, in this case, a singular integral equation of Cauchy type which can be solved in closed form by using well-known techniques.[18,19] The limiting form of Eq. (2.26) is obtained by using the Plemelj formulas in the form

$$\lim_{\epsilon \to +0} \frac{1}{x \pm i\epsilon} = \mp \pi i \delta_*(x). \tag{4.1}$$

Equation (2.25) thus becomes

$$\epsilon^-(\nu - w)F_{\mathbf{k}}(\mathbf{v}; w) = \pi i\, d_{\mathbf{k}}(\mathbf{v}) \int d\mathbf{v}_1\, \delta_-(\mathbf{k}\cdot\mathbf{v} - kw - \mathbf{k}\cdot\mathbf{v}_1) \times F_{-\mathbf{k}}(\mathbf{v}_1; w) + q_{\mathbf{k}}(\mathbf{v}; w). \tag{4.2}$$

This equation has exactly the same form as the equation derived by Guernsey[7] using a different method (it can be derived from his Eq. (21) by the substitution $\mathbf{k} \to -\mathbf{k}$, $\mathbf{k}' \to \mathbf{k}$). Although he also considers inhomogeneous systems[20] his method is restricted to systems close to equilibrium. For this reason, the coefficients ϵ^-, $d_{\mathbf{k}}$, and $q_{\mathbf{k}}$ in his equation are expressed in terms of the Maxwell distribution of velocities. Our present method does not rely on such an assumption, and therefore the coefficients are functionals of the time-dependent distribution of velocities (which is however retarded in time). The possibility of relaxing the assumption of small departures from equilibrium is, in our opinion, an important generalization. Only with such a method can we envisage the possibility of studying unstable systems which are intrinsically very far from thermal equilibrium.

We now note that Guernsey's method of solution (Sec. III of reference 7) depends crucially on the fact that $\epsilon^-(\nu)$ [or $2\,\mathfrak{D}^-(u)$ in his notation] has no zeros in the lower half-plane [see the passage from (52) to (53), and from (55) to (58) in reference 7]. Hence, for *stable systems*, Guernsey's method of solution can be taken over directly in our case and results in the following formula[21]

$$F_{\mathbf{k}}(\mathbf{v}; w) = \frac{q_{\mathbf{k}}(\mathbf{v}; w)}{\epsilon^-(\nu - w)} + \pi i \frac{d_{\mathbf{k}}(\mathbf{v})}{k} \int_{-\infty}^{\infty} d\nu_1\, \delta_-(\nu - \nu_1) \times \frac{[\bar{q}_{\mathbf{k}}(\nu_1; w)]_+ - [\bar{q}_{-\mathbf{k}}(-\nu_1 + w; w)]_-}{\epsilon^+(\nu_1)\epsilon^-(\nu_1 - w)}. \tag{4.3}$$

In this and subsequent equations, ν denotes the component of $\mathbf{v}$ parallel to $\mathbf{k}$:

$$\nu = \mathbf{k}\cdot\mathbf{v}/k.$$

The barred functions have been defined by (3.1). The functions bearing a subscript + or − are defined as follows: Given an arbitrary integrable

[18] N. I. Muskhelishvili, *Singular Integral Equations* (P. Noordhof N. V., Groningen, Holland, 1953).

[19] S. G. Mikhlin, *Singular Integral Equations*, Transl. Am. Math. Soc., No. 24 (1950).

[20] An extension of the present method to inhomogeneous systems will be published soon.

[21] An alternative method of solution, simpler than Guernsey's, can be found in reference 2, Appendix 8. It is based on the properties of the Van Kampen Case eigenfunctions of the Vlassov equations.

function $f(v)$ of the real variable v, it can be split uniquely into a difference of a *"plus function"* and a *"minus function"*:

$$f(v) = f_+(v) - f_-(v), \tag{4.4}$$

having the following properties: A plus function has a regular analytic continuation into the upper half-plane S_+; correspondingly, the analytic continuation of $f_-(v)$ into the lower half-plane has no singularities there. From our discussion of the previous section it follows that the Plemelj formula (3.7) provides precisely this decomposition. Hence,

$$f_\pm(v) = \pm\frac{1}{2}\int dv_1\ \delta_\pm(v - v_1)f(v_1). \tag{4.5}$$

Using the definition (2.28) as well as (4.1), it is easily shown that Eq. (4.3) can be rewritten as follows[22] [see Fig. 3(a)]:

$$F_k(\mathbf{v}; w) = \frac{q_k(\mathbf{v}; w)}{\epsilon^-(v - w)} + \frac{\omega_p^2}{4\pi k^2}\frac{d_k(\mathbf{v})}{k}\int dv_1\ \delta_-(v - v_1) \times \frac{\bar\varphi_-'(v_1 - w)\bar\varphi_+(v_1) - \bar\varphi_-(v_1 - w)\bar\varphi_+'(v_1)}{\epsilon^+(v_1)\epsilon^-(v_1 - w)}$$

(stable; w real). (4.6)

In this formula it is readily verified that $F_k(\mathbf{v}; w)$ has no singularities (in w) in the upper half-plane. Hence, as stated at the beginning of this section, Eq. (4.6) as it stands is also, for $w \in S_+$, a solution of the original Eq. (2.25).

5. SOLUTION OF THE INTEGRAL EQUATION (2.25) IN THE UNSTABLE CASE

We are now in a position to see why the previous method fails in the unstable case. The main idea in the previous paragraph was to replace the solution of the original Eq. (2.25) by the solution of the simpler Eq. (4.2), obtained for real w. Suppose now that the system is unstable, i.e. $\zeta_+ \in S_+$ and $\zeta_- \in S_-$. The auxiliary Eq. (4.2) can still be solved exactly. The solution will actually no longer be (4.3), because the index of the singular integral equations appearing in Guernsey's paper changes. However, there exist standard methods[18,19] for solving the problem in closed form even in this case, and it turns out that one obtains the stable terms (4.3) plus two extra terms. The main point of the discussion is the fact that $F_k(\mathbf{v}; w)$ has now a pole in $w = v - \zeta_-$, and $F_{-k}(\mathbf{v}; w)$ has therefore a pole in $w = v + \zeta_+$, both located in the upper half-plane. This can be seen already in the first term of (4.3). It will be easily understood that the solution of Eq. (4.2) is no longer a solution of (2.26). Indeed, consider the first term in the r. h. s. of the latter equation, which contains the integral

$$\int_{-\infty}^{\infty} dv_1\,\frac{\bar F_{-k}(-v_1; w)}{v_1 - v + w}, \quad \mathrm{Im}\ w > \gamma_0.$$

This Cauchy integral has a cut on the real axis $\mathrm{Im}\ w = 0$. But if F_{-k} has a pole in $w = v + \zeta_+$, this integral contains at least some terms of the form

$$\int_{-\infty}^{\infty} dv_1\,\frac{\psi(v_1; w)}{(v_1 - v + w)(v_1 - w + \zeta_+)}, \quad \mathrm{Im}\ w > \gamma_0.$$

Such an integral has a cut on the line $\mathrm{Im}\ w = \gamma_0 > 0$. Hence, by replacing the previous integral simply by

$$\pi i\int_{-\infty}^{\infty} dv_1\,\frac{\psi(v_1; w)}{(v_1 - w + \zeta_+)}\,\delta_-(v - w - v_1), \quad w \text{ real},$$

for real w we are no longer on the same branch of the function as given by the previous expression for $\mathrm{Im}\ w > \gamma_0$ (see the discussion in Sec. 3). The auxiliary Eq. (4.2) is therefore *not* the analytic continuation of Eq. (2.25). Hence the analytic continuation of the solution (4.6) for $\mathrm{Im}\ w > \gamma_0$ does not satisfy the original Eq. (2.25) in the case of unstable plasmas, as it did in the stable case (where the cut $\mathrm{Im}\ w = \gamma_0$ lies *below* the real axis.)

However the knowledge of the stable solution

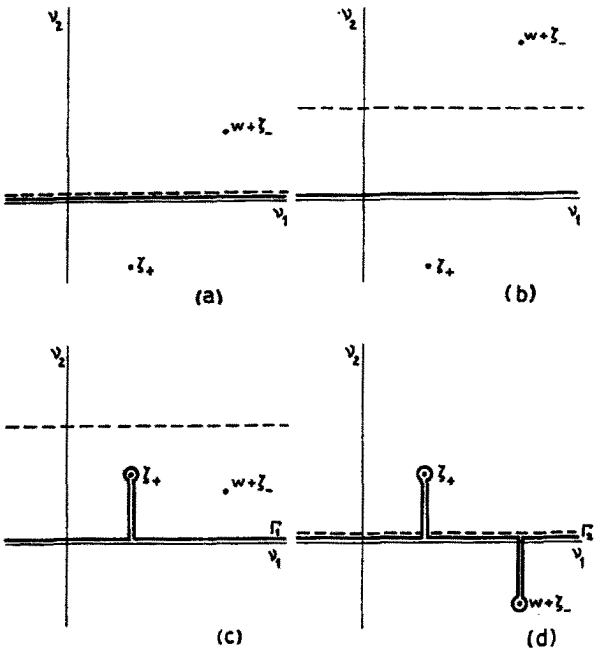

FIG. 3. Contours of integration for $F_k(\mathbf{v}; w)$ in various situations. (a) stable plasma, w real; (b) stable plasma, $w \in S_+$; (c) unstable plasma, $w \in S_+$; (d) unstable plasma, w real.

[22] In handling double principal-part integrals, the Poincaré–Bertrand theorem[18,19] must be used when the order of integrations is changed.

enables us to obtain the solution in the unstable case too, by appropriate deformations of the contours of integration. Consider again the stable case, and let w be on the original Bromwich contour C, above all the singularities of F: Eq. (4.6) provides the solution in that case. Let us introduce the following abbreviation:

$$\bar{R}(v; w) = i\frac{\omega_p^2}{4\pi^2 ck^2} \times \frac{\bar{\varphi}_-'(v-w)\bar{\varphi}_+(v) - \bar{\varphi}_-(v-w)\bar{\varphi}_+'(v)}{\sigma^+(v)\sigma^-(v-w)}. \quad (5.1)$$

Then the solution is [see Fig. 3(b)]

$$F_{\mathbf{k}}(\mathbf{v}; w) = \frac{q_{\mathbf{k}}(\mathbf{v}; w)}{\epsilon^-(v-w)} + \pi i\frac{d_{\mathbf{k}}(\mathbf{v})}{k} \times \int_{-\infty}^{\infty} dv_1\ \delta_-(v-v_1)\frac{\bar{R}(v_1; w)}{(v_1-\zeta_+)(v_1-w-\zeta_-)}$$

(stable; $\operatorname{Im} w > |\gamma_0|$). (5.2)

Let now ζ_+ move into the upper half-plane and ζ_- into the lower half-plane, keeping w constant. As long as w is sufficiently far up in S_+, Eq. (2.25) changes continuously in this process. Hence in the unstable case, the solution of Eq. (2.25) will be the analytic continuation of (5.2) for $\zeta_+ \in S_+$. The latter is obtained by deforming the contour of the v_1 integration as shown in Fig. 3(c):

$$F_{\mathbf{k}}(\mathbf{v}; w) = \frac{q_{\mathbf{k}}(\mathbf{v}; w)}{\epsilon^-(v-w)} + \frac{\pi i}{k} d_{\mathbf{k}}(\mathbf{v}) \times \int_{\Gamma_1} dv_1\ \delta_-(v-v_1)\frac{\bar{R}(v_1; w)}{(v_1-\zeta_+)(v_1-w-\zeta_-)}$$

(unstable; $\operatorname{Im} w > \gamma_0$). (5.3)

This equation is the solution to our problem. For further calculations, it is usually convenient to have expressions in terms of real w. We therefore move down the contour C and bring it on the upper edge of the real axis, keeping now ζ_+ and ζ_- constant. The pole $w + \zeta_-$ thus moves into the lower half-plane, and the analytical continuation of (5.3) is obtained by deforming the contour of integration into the contour Γ_2 shown in Fig. 3(d). The result, expressed in terms of an integral over the real axis, is therefore

$$F_{\mathbf{k}}(\mathbf{v}; w) = \frac{q_{\mathbf{k}}(\mathbf{v}; w)}{\epsilon^-(v-w)} + \frac{2\pi i}{k} d_{\mathbf{k}}(\mathbf{v}) \times \left\{\frac{1}{2}\int_{-\infty}^{\infty} dv_1\ \delta_-(v-v_1)\frac{\bar{R}(v_1; w)}{(v_1-\zeta_+)(v_1-w-\zeta_-)} + \frac{\bar{R}(\zeta_+; w)}{(w+\zeta_--\zeta_+)(v-\zeta_+)} + \frac{\bar{R}(w+\zeta_-; w)}{(w+\zeta_--\zeta_+)(v-\zeta_--w)}\right\}$$

(unstable; w real). (5.4)

This is the final form of the solution. Our initial statement can now be verified *a posteriori:* (5.4) is not a solution of Eq. (4.2).[23]

6. KINETIC EQUATIONS FOR WEAKLY UNSTABLE AND FOR WEAKLY STABLE PLASMAS

Equation (5.4), substituted into Eqs. (2.24) and (2.23) provides the explicit form of the kinetic equation for unstable systems. It is an extremely complicated non-Markoffian equation. [Remember that all functions $\varphi(\mathbf{v})$ occuring in (5.4) are evaluated at time $t - \tau$.] The equation can, however, be appreciably simplified in the case of *"weakly unstable" plasmas*, a concept which will now be defined. Assume that the imaginary part of the "unstable" zero of the dielectric constant γ_0 is much smaller than the imaginary part of the stable zero closest to the real axis, which we denote by γ_m; the latter has been assumed to be of order $\omega_p/\bar{k}$, i.e. $\bar{k}\gamma_m$ measures the short time scale of the plasma:

$$\gamma_0 \ll |\gamma_m| \approx \omega_p/\bar{k}. \quad (6.1)$$

We also assume (this must be verified in the result) that if (6.1) is satisfied at the initial time, it remains true at all later times, as long as $\gamma_0 > 0$.

We can now distinguish in (2.24) a slow process (described by the residues at the unstable poles) and rapidly damped transient processes. Hence, if we are interested in times much longer than ω_p^{-1} we can drop the latter contributions. Moreover, as was shown in Sec. 2, we can use the Markoffian approximation (2.21), by setting $\tau = 0$ in all factors $\varphi(t-\tau)$ appearing in the function $F_{\mathbf{k}}$. Hence (2.24) becomes

$$\mathfrak{F}_{\mathbf{k}}(\mathbf{v}; t) = -\frac{1}{2\pi i}\int_c dw\,\frac{e^{-ikwt}}{w} F_{\mathbf{k}}(\mathbf{v}; w). \quad (6.2)$$

However, we are now interested in times which can be of the order of $(\bar{k}\gamma_0)^{-1}$ which is much longer than the period of a plasma oscillation. We must therefore take into account not only the residue in $0+$, but also the residues at the unstable poles.

In order to evaluate these residues, we first inte-

[23] An alternative, completely independent method of summation of both stable and unstable rings can be found in reference 2, Appendix 10. It is based on a factorization theorem proven by P. Résibois, Phys. Fluids 6, 817 (1963).

grate explicitly over ν_1 the second term in Eq. (5.4). This is most easily done by using the decomposition of $\bar{R}(\nu_1; w)$ (regarded as a function of ν_1) into plus and minus parts according to the Plemelj formula (4.4). Using also the expression of $\delta_-(x)$ given by (4.1) and closing the contour of integration, we obtain

$$F_{\mathbf{k}}(\mathbf{v}; w) = \frac{q_{\mathbf{k}}(\mathbf{v}; w)}{\epsilon^-(\nu - w)} + \frac{2\pi i}{k} d_{\mathbf{k}}(\mathbf{v}) \left\{ -\frac{\bar{R}_-(\nu; w)}{(\nu - w - \zeta_-)(\nu - \zeta_+)} - \frac{\bar{R}_-(\zeta_+; w)}{(w + \zeta_- - \zeta_+)(\nu - \zeta_+)} + \frac{\bar{R}_+(w + \zeta_-; w)}{(w + \zeta_- - \zeta_+)(\nu - \zeta_- - w)} \right\}. \tag{6.3}$$

Substituting this expression into (6.2), the Laplace transform is easily calculated by the method of residues. Noting that $q_{\mathbf{k}}(\mathbf{v}; w)$ and $\bar{R}_-$ have poles far down in the lower half-plane, thus giving rapidly damped terms which can be neglected, we obtain

$$\mathfrak{F}_{\mathbf{k}}(\mathbf{v}; t) = \frac{q_{\mathbf{k}}(\mathbf{v}; 0)}{\epsilon^-(\nu)} + \frac{2\pi i}{k} d_{\mathbf{k}}(\mathbf{v}) \left\{ -\frac{\bar{R}_-(\nu; 0)}{(\nu - \zeta_+)(\nu - \zeta_-)} + \frac{1}{2i\gamma_0} \left[\frac{\bar{R}_-(\zeta_+; 0)}{\nu - \zeta_+} - \frac{\bar{R}_+(\zeta_-; 0)}{\nu - \zeta_-} \right] \right\} + \frac{e^{2k\gamma_0 t}}{2i\gamma_0} \frac{2\pi i}{k} d_{\mathbf{k}}(\mathbf{v}) \frac{\bar{R}(\zeta_+; \zeta_+ - \zeta_-)}{\nu - \zeta_+} + \frac{e^{-ik(\nu - \zeta_-)t}}{\nu - \zeta_-} \left\{ -\frac{q_{\mathbf{k}}(\mathbf{v}; \nu - \zeta_-)}{\sigma^-(\zeta_-)} - \frac{2\pi i}{k} d_{\mathbf{k}}(\mathbf{v}) \frac{\bar{R}(\nu; \nu - \zeta_-)}{\nu - \zeta_+} \right\}. \tag{6.4}$$

This expression substituted into (2.23) provides the general kinetic equation for weakly unstable systems. It consists of a term independent of time [except through $\varphi(\mathbf{v}; t)$], an exponentially growing term, and an oscillating term which is exponentially amplified.

Before discussing this equation further, we note the following important fact: After stabilization of the plasma, the formerly unstable zero ζ_+ moves into the lower half-plane. However it remains for a certain time much closer to the real axis than all the other zeros of $\epsilon^+(w)$. Hence, if we want an asymptotic description of such a *"weakly stable" plasma* to the same degree of precision as (6.4), we must retain in (6.2) the residues at the poles related to ζ_+, in addition to the residue in $w = 0$ which is considered in "normal" stable plasmas. But in the stable case, $F_{\mathbf{k}}(\mathbf{v}; w)$, regarded as a function of the parameter ζ_+, is the analytic continuation of this same function for the unstable case. Hence the expression for $\mathfrak{F}_{\mathbf{k}}(\mathbf{v}; t)$ has the same analytic form (6.4) in both cases, as can also be verified directly from (4.6). Of course, in the stable case, $\zeta_+ \in S_-$, $\zeta_- \in S_+$, $\gamma_0 < 0$.

We now proceed to simplify the kinetic equation. We first note that

$$-\frac{\bar{R}_-(\nu; 0)}{(\nu - \zeta_+)(\nu - \zeta_-)} + \frac{1}{2i\gamma_0} \left\{ \frac{\bar{R}_-(\zeta_+; 0)}{\nu - \zeta_+} - \frac{\bar{R}_+(\zeta_-; 0)}{\nu - \zeta_-} \right\} = \frac{1}{2} \int d\nu_1 \, \delta_-(\nu - \nu_1) \frac{\bar{R}(\nu_1; 0)}{(\nu_1 - \zeta_+)(\nu_1 - \zeta_-)} - \frac{1}{2i\gamma_0} \left\{ \frac{\bar{R}(\zeta_+; 0)}{\nu - \zeta_+} + \frac{\bar{R}(\zeta_-; 0)}{\nu - \zeta_-} \right\}.$$

It is easily shown that the bracketted term in the r. h. s. is an odd function of the wave-vector $\mathbf{k}$; hence it does not contribute to the kinetic equation (2.23).

Consider now the last term in (6.4), proportional to $\exp[-ik(\nu - \zeta_-)t]$. There is a certain inconsistency in retaining rapid oscillations in an asymptotic equation valid for times $t \gg \omega_p^{-1}$. We must in some way smooth out these rapid oscillations, and retain only their slow systematic growth. This can be achieved by noting that $\varphi(\mathbf{v}; t)$ is actually a distribution in the sense of L. Schwartz: the physically relevant quantities are integrals of products of φ and of some function of $\mathbf{v}$. Let $U(\mathbf{v})$ be such a function, which we can assume to be an entire function (in all physical applications $U(\mathbf{v})$ is a polynomial). Let us call ν and $\mathbf{v}_\perp$, respectively, the components of $\mathbf{v}$ parallel and perpendicular to $\mathbf{k}$. Whenever required by clarity, we will write functions $f(\mathbf{v})$ of the vector $\mathbf{v}$ in the forms $f(\mathbf{v}_\perp, \nu)$. Multiplying now both sides of the kinetic equation (2.23) by the test function $U(\mathbf{v}_\perp, \nu)$ and integrating over $\mathbf{v}$, the last term in (6.4) gives a contribution of the following form to $\partial_t \langle U \rangle$:

$$\int_{-\infty}^{\infty} d\nu \frac{e^{-ik(\nu - \zeta_-)t}}{\nu - \zeta_-} \left\{ f_1(\nu) + \frac{f_2(\nu)}{\nu - \zeta_+} \right\}. \tag{6.5}$$

For positive times the contour of integration is closed in the lower half-plane. In the *unstable case*, the pole $\nu = \zeta_-$ is thus within the contour, whereas $\nu = \zeta_+$ is outside. The functions $f_1(\nu)$ and $f_2(\nu)$ have poles far down in the lower half-plane, and thus give rapidly damped residues which are neglected in order to be consistent with our approximations. This asymptotic result of cutting all the poles but

the one in ζ_- is achieved by replacing the integrand in (6.5) by

$$\frac{1}{\nu - \zeta_-}\left\{f_1(\zeta_-) + \frac{f_2(\zeta_-)}{\zeta_- - \zeta_+}\right\}.$$

Consider now the *stable case*. The pole $\nu = \zeta_-$ is now outside the contour whereas $\nu = \zeta_+$ has moved inside. An argument similar to the previous one shows that the asymptotic form of (6.5) is obtained by replacing the integrand by

$$\frac{1}{\nu - \zeta_+} e^{-ik(\zeta_+ - \zeta_-)t} \frac{f_2(\zeta_+)}{\zeta_+ - \zeta_-}.$$

As a result of this discussion, the consistent asymptotic forms of the kinetic equation (2.23) are

$$\partial_t\varphi(\mathbf{v};t) = -\omega_p^2 \int d\mathbf{k}\, k^{-2} i\mathbf{k}\cdot\partial \left\{\frac{q_\mathbf{k}(\mathbf{v};0)}{\epsilon^-(\nu)}\right.$$
$$+ \frac{\pi i}{k} d_\mathbf{k}(\mathbf{v}) \int d\nu_1\, \delta_-(\nu - \nu_1) \frac{\bar{R}(\nu_1;0)}{(\nu_1 - \zeta_+)(\nu_1 - \zeta_-)}$$
$$+ \frac{1}{\nu - \zeta_-}\left[-\frac{q_\mathbf{k}(\mathbf{v}_\perp, \zeta_-;0)}{\sigma^-(\zeta_-)}\right.$$
$$\left.\left. + \frac{2\pi i}{k} d_\mathbf{k}(\mathbf{v}_\perp, \zeta_-) \frac{\bar{R}(\zeta_-;0)}{2i\gamma_0}\right]\right.$$
$$\left. + \frac{1}{\nu - \zeta_+} \frac{2\pi i}{k} d_\mathbf{k}(\mathbf{v}) e^{2k\gamma_0 t} \frac{\bar{R}(\zeta_+;\zeta_+ - \zeta_-)}{2i\gamma_0}\right\}$$
(unstable), (6.6)

and

$$\partial_t\varphi(\mathbf{v};t) = -\omega_p^2 \int d\mathbf{k}\, k^{-2} i\mathbf{k}\cdot\partial \left\{\frac{q_\mathbf{k}(\mathbf{v};0)}{\epsilon^-(\nu)}\right.$$
$$+ \frac{\pi i}{k} d_\mathbf{k}(\mathbf{v}) \int d\nu_1\, \delta_-(\nu - \nu_1)$$
$$\times \frac{\bar{R}(\nu_1;0)}{(\nu_1 - \zeta_+)(\nu_1 - \zeta_-)} + \frac{2\pi i}{k} \frac{1}{\nu - \zeta_+}$$
$$\left.\times \frac{e^{2k\gamma_0 t}\bar{R}(\zeta_+;\zeta_+ - \zeta_-)}{2i\gamma_0} [d_\mathbf{k}(\mathbf{v}) - d_\mathbf{k}(\mathbf{v}_\perp, \zeta_+)]\right\}$$
(stable). (6.7)

Let us now write these expressions more explicitly. We first note that the residue in $w = 0$ is exactly identical in form to the function $F_\mathbf{k}(\mathbf{v}; 0+)$ appearing in (2.22); hence it gives precisely the "normal collision term" $\mathcal{C}\{\varphi\}$ defined in Eq. (1.2). The other terms, rewritten explicitly by using (5.1) give the following expressions:

$$\partial_t\varphi(\mathbf{v};t) = \mathcal{C}\{\varphi\} + \mathcal{J}\{\varphi;t\}, \tag{6.8}$$

with

$$\mathcal{J}\{\varphi;t\} = \frac{4e^4c}{m^2} \int d\mathbf{k}\, k^{-4}\mathbf{k}\cdot\partial \left\{-\frac{1}{2\pi i}\frac{k^2}{\omega_p^2}\frac{\varphi(\mathbf{v}_\perp, \zeta_-)}{(\nu - \zeta_-)\sigma^-(\zeta_-)}\right.$$
$$+ \frac{\bar{\varphi}(\zeta_-)}{2ik\gamma_0(\nu - \zeta_-)\sigma^+(\zeta_+)\sigma^-(\zeta_-)} \mathbf{k}\cdot\partial\varphi(\mathbf{v}_\perp, \zeta_-)$$
$$\left. + \frac{D(t)}{\mathbf{k}\cdot\mathbf{v} - k\zeta_+} \mathbf{k}\cdot\partial\varphi(\mathbf{v})\right\}$$
(unstable), (6.9)

and

$$\mathcal{J}\{\varphi;t\} = \frac{4e^4c}{m^2} \int d\mathbf{k}\, k^{-4}\mathbf{k}\cdot\partial \frac{D(t)}{\mathbf{k}\cdot\mathbf{v} - k\zeta_+}$$
$$\times \mathbf{k}\cdot\partial\{\varphi(\mathbf{v}) - \varphi(\mathbf{v}_\perp, \zeta_+)\} \quad \text{(stable)}. \tag{6.10}$$

The time-dependent coefficient $D(t)$ is defined in both cases by

$$D(t) = \frac{e^{2k\gamma_0 t}}{2i\gamma_0} \frac{\bar{\varphi}_+(\zeta_+) - \bar{\varphi}_-(\zeta_-)}{\sigma^+(\zeta_+)\sigma^-(\zeta_-)}. \tag{6.11}$$

7. MECHANISM OF EVOLUTION

The discussion of the detailed mechanism of evolution of an initially unstable plasma has to await further investigation of the properties of Eq. (6.8), in particular, the study of some simple examples. However, we may draw some qualitative conclusions from the form of this equation.

Equation (6.8) has the general structure of a Fokker–Planck equation, i.e. it contains a friction term and a diffusion term. A characteristic feature of the equation is the exponential time dependence of the diffusion and friction coefficients. The latter are, moreover, functionals of the distribution function; hence their form changes as the latter function evolves in time. The zero ζ_+ of the dielectric constant is itself a functional of $\varphi(\mathbf{v})$. Hence the overall process is a very complex nonlinear friction and diffusion phenomenon in velocity space.

We keep in mind[10–12] the fact that an unstable plasma is characterized by a velocity distribution with two humps, which are sufficiently widely separated. Both friction and diffusion are stabilizing agents: the first brings the two maxima closer together, the second one broadens them. As a consequence, the zero ζ_+ will move down until the critical separation of the maxima is attained. At this point, the zero ζ_+ attains the real value w_0 (neutral stability) and the plasma is stabilized. The exponential time dependence of the friction and diffusion coefficients make this process more effective the more unstable the plasma (i.e. the larger γ_0). The transition through the neutral point will be discussed below. The evolution of the weakly stable plasma resulting from this process then continues towards thermal

equilibrium under the combined action of the normal collision term and of the extra term of Eq. (6.10). The latter becomes less and less effective as the plasma becomes more and more stable (i.e. as ζ_+ moves down in S_-). In the final stage, this term becomes negligible and the plasma evolves towards equilibrium under the action of the normal collision term alone.

We now study more closely the transition through the neutral stability and show that the two forms of Eq. (6.8) go continuously into one another. The discussion is most clearly performed on Eqs. (6.6) and (6.7). We first note that, although the first bracketted term in the r. h. s. has the same form in both equations, its behavior is different in stable and unstable cases. Indeed, introducing again a test function $U(\mathbf{v})$ and integrating over $\mathbf{v}$, as was done in Sec. 6, we obtain schematically the following contributions:

(a) $q_{\mathbf{k}}(\mathbf{v}_\perp, \zeta_-; 0)/\sigma^-(\zeta_-)$ + Sum of residues at poles of $q_{\mathbf{k}}$ in S_- (unstable case),

(b) Sum of residues at poles of $q_{\mathbf{k}}$ in S_- (stable case).

We see that there is an extra term in the unstable case. But the contribution of this extra term is exactly canceled by the third term in (6.6). Hence we may write the following equivalence relation in the sense of distributions:

$$\left\{\frac{q_{\mathbf{k}}(\mathbf{v}_\perp, \nu; 0)}{\epsilon^-(\nu)} - \frac{q_{\mathbf{k}}(\mathbf{v}_\perp, \zeta_-; 0)}{(\nu - \zeta_-)\sigma^-(\zeta_-)}\right\}_{\text{unstable}} \sim \left\{\frac{q_{\mathbf{k}}(\mathbf{v}_\perp, \nu; 0)}{\epsilon^-(\nu)}\right\}_{\text{stable}}.$$

It is easily seen that the integral terms in both equations tend towards the same limit as $\gamma_0 \to 0$. Consider now the last two terms in Eq. (6.8). They have the following limit as $\gamma_0 \to 0$:

$$\frac{2\pi i}{k}\{-\pi i\delta_+(\nu - w_0)\, d_{\mathbf{k}}(\mathbf{v}_\perp, w_0) + \pi i\delta_-(\nu - w_0) \times d_{\mathbf{k}}(\mathbf{v}_\perp, \nu)(1 + 2k\gamma_0 t + \cdots)\}\frac{\bar{R}(w_0; 0)}{2i\gamma_0}.$$

It is easily verified that, for reasons of parity in $\mathbf{k}$, only the δ part of the $\delta_\pm$ functions contributes to the kinetic equation. Hence the divergent parts cancel each other and there remains a term proportional to $d_{\mathbf{k}}(\mathbf{v}_\perp, w_0)$. This term, however, vanishes because $\bar{\varphi}'(w_0)$ (for real w_0) is the imaginary part of $\epsilon^+(w_0)$, and this implies in turn that $\mathbf{k}\cdot\partial\,\varphi(\mathbf{v}_\perp, w_0) = 0$. The same argument shows that the exponential term in (6.7) also vanishes at neutral stability. This achieves the proof that the two forms of the kinetic equation go over continuously into one another.

In conclusion, we may point out that Eq. (6.8) deserves special interest from the point of view of the general nonequilibrium statistical mechanics. It is a Markoffian equation, in which however the non-Markoffian contributions leave a trace even for long times. Such a situation is not possible in the case of ordinary gases with short-range interactions (where the poles of the resolvent are fixed only by the interaction potential). In a plasma, however, the position of the poles of the effective resolvent depends on the form of the velocity distribution, and can move quite close to the real axis or even cross it, as it does in the present problem.

We intend to study in subsequent papers the further details of the mechanism of evolution of unstable plasmas, as well as other aspects and extensions of the present theory (density correlations, inhomogeneous systems, etc.).

ACKNOWLEDGMENTS

We wish to express our thanks to Professor I. Prigogine for the interest he took in this work. We would like to acknowledge numerous fruitful discussions with Dr. B. Abraham and Dr. A. Kuszell. The first stages of this work were performed during the military service of the author at the École Royale Militaire in Brussels; we wish to thank Professor E. Thomas and Dr. P. Vandenplas for their kind hospitality at the Department of Physics.

This work has been partially supported by the United States Air Force under Contract No. AF 61(052)-179, and monitored by the European Office, Office of Aerospace Research.

Regularity of the T Matrix in the Case of Dirac Potential Scattering*

ANTONIO PAGNAMENTA†
Department of Physics and Astronomy, University of Maryland, College Park, Maryland
(Received 26 December 1962)

Applying Hunziker's method to the case of Dirac potential scattering, we prove analyticity of the T matrix in energy and momentum transfer. Our conditions on the potential are somewhat weaker, and the domains of analyticity contain the known ones as special cases.

I. INTRODUCTION

We discuss the analytic properties of the scattering amplitude, or the T matrix, for the scattering of a positive energy Dirac particle from a central potential.

Following a suggestion by Hunziker,[1] that the method he developed to discuss the Schrödinger scattering amplitude be also applicable in the Dirac case, we use essentially his method. To derive the regularity properties of the T matrix, we use directly the general functional analytical properties of the integral equation and its kernel. No series development is used. Since the only known discussion of this case is the one of Khuri and Treiman,[2] we adopt their notation basically.

In the second section, we consider the scattering equation as a functional equation in the space of the continuous and bounded functions. We then establish conditions on the potential, in order that the integral operator KV be bounded.

We can show that $KV(k)$ is also completely continuous for k in the upper half-plane. Therefore, the Fredholm alternative applies. From this we get all the necessary information on the wavefunction that we need to discuss the transition matrix.

We show that, for a wide class of potentials, the T matrix is regular in the energy and momentum transfer variables, each in an established domain. The restrictions on the potential are naturally stronger than in the Schrödinger case, but less restrictive than in reference 2. Since we have to deal here not only with a single equation but with a set of coupled differential equations, the discussions are somewhat more involved than for the Schrödinger equation. In order to preserve continuity, we have not omitted some short proofs which are given already in Hunziker's work.

II. SOME PROPERTIES OF THE DIRAC SCATTERING EQUATION

A. The Dirac Scattering Equation

Consider the scattering of a Dirac particle of mass m, total energy E, in a central field $V(r)$. Let us write the Dirac equation in the following form ($\hbar = C = 1$):

$$(E - H_D)\psi = V\psi. \tag{1}$$

Here ψ is a four-component spinor wavefunction and

$$H_D = -i\boldsymbol{\alpha}\cdot\nabla + \beta m$$

is the free-particle Hamiltonian. $\boldsymbol{\alpha}$ and β are the usual Dirac matrices. Denote by ϕ plane wave solutions of the free-particle equation

$$(E - H_D)\phi = 0,$$

which are characterized by the energy and momentum eigenvalues, as well as by their spin, namely

$$\phi = u(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}}$$

$u(\mathbf{k})$ is a four-spinor normalized to

$$u^\dagger u = 1,$$

with k real and $k = |\mathbf{k}| = (E^2 - m^2)^{\frac{1}{2}}$, $E \geq m$.

To describe the scattering of a particle with a certain initial momentum and spin (described by ϕ), we look for a solution of (1) that has the asymptotic behavior of a plane wave plus an outgoing spherical wave:

$$\psi \underset{r\to\infty}{\longrightarrow} [e^{i\mathbf{k}\cdot\mathbf{x}} + (e^{ikr}/r)f_{op}]u,$$

where f_{op} is an operator in spinor space. This boundary condition is automatically incorporated in the integral equation formulation of (1) which in operator notation can be written

$$\psi = \phi + \lim_{\epsilon\to+0}(E - H_D + i\epsilon)^{-1}V\psi. \tag{2}$$

† This work is based in part on the author's MS thesis, submitted at Eidgenössische Technische Hochschule, Switzerland, (September, 1961).

* Supported by the U. S. Air Force.

[1] W. Hunziker, Helv. Phys. Acta **34**, 593 (1961).

[2] N. N. Khuri and S. B. Treiman, Phys. Rev. **109**, 198 (1958).

The formal solution of this equation is

$$\psi = \phi + \lim_{\epsilon \to +0} (E - H_D - V + i\epsilon)^{-1} V\phi.$$

We only discuss the scattering of particles ($E \geq m$) explicitly, the connection with antiparticles is given in reference 2.

Using the outgoing wave Green's function

$$G_0(\mathbf{x}, \mathbf{y}) = (-1/4\pi)(|\mathbf{x} - \mathbf{y}|)^{-1} \exp [ik\,|\mathbf{x} - \mathbf{y}|],$$

we can write (2)

$$\psi_\mu(\mathbf{x}) = \phi_\mu(\mathbf{x}) - \frac{1}{4\pi} \int (E - i\boldsymbol{\alpha}\cdot\nabla_{\mathbf{x}} + \beta m)_{\mu\nu}$$

$$\times (|\mathbf{x} - \mathbf{y}|)^{-1} \exp [ik\,|\mathbf{x} - \mathbf{y}|] V(\mathbf{y})\psi_\nu(\mathbf{y})\, d^3y. \quad (3)$$

Here we have just for once indicated the spinor indices (Greek subscripts). We shall drop them from now on, and all our further notation has to be considered as a matrix notation, with the straightforward implication on norms and absolute values to be taken for each matrix element and spinor component.

Our method allows us to carry out the discussion of the functional analytical properties of the Dirac scattering functions directly on Eq. (3). We do not need any further iterations or the second-order equation. Let us introduce in (3) the following notation:

$$\psi(\mathbf{x}) = u(\mathbf{k}) \exp [i\mathbf{k}\cdot\mathbf{x}] + \int d^3y\, K(\mathbf{x}, \mathbf{y}) V(\mathbf{y})\psi(\mathbf{y}), \quad (4)$$

with

$$K(\mathbf{x}, \mathbf{y}) = [E + \beta m - i\alpha_k(\partial/\partial x_k)] G_0(\mathbf{x}, \mathbf{y}).$$

In this form one could extract the bracket from the integral. Since integral operators are much handier, however, than differential operators, we leave it and do also the differentiation on $G_0(\mathbf{x}, \mathbf{y})$. Thus we get for the kernel K, which we split for later use into two characteristical terms,

$$K(\mathbf{x}, \mathbf{y}) = \left(E + \beta m + k\alpha_i \frac{x_i - y_i}{|\mathbf{x} - \mathbf{y}|}\right) G_0 + i\alpha_i \frac{x_i - y_i}{|\mathbf{x} - \mathbf{y}|} \cdot \frac{G_0}{|\mathbf{x} - \mathbf{y}|}.$$

In the operator notation, (4) becomes

$$\psi = \phi + KV\psi. \quad (5)$$

B. Conditions on the Potential

In this section we establish the conditions on the potential, in order that the operator KV has a finite norm. We show that, for the complex k, KV is completely continuous in Im $k > 0$. These are the conditions on an integral equation of the second kind that the Fredholm alternative applies.

Define the norm of ψ by

$$||\psi|| = \sup_{\mathbf{x}, i} |\psi_i(\mathbf{x})|. \quad (6)$$

The $\mathbf{x}$ means that we have to take the supremum of the absolute value for all $\mathbf{x}$, and the i is set to indicate that this has to be done for all spinor components.

At first we shall look for a solution of (5) in the space of C which contains all continuous and bounded functions. C is a complete space, normed with (6); therefore, we are working with elements of a Banach space.

Define the norm of an operator O by

$$||O|| = \sup ||O\psi|| \quad \psi \in C \quad \text{and} \quad ||\psi|| \leq 1.$$

We prove in Appendix I that this norm is finite for our operator KV in (5) under the following assumptions on the potential:

$$\int_0^M |V(x)|\, dx < \infty, \quad \text{I.}$$

$$\int_M^\infty x\, |V(x)|\, dx < \infty, \quad \text{II.}$$

$$\int_0^\infty x^\epsilon\, |V(x)|^{1+\epsilon}\, dx < \infty. \quad \text{III.}$$

M is an arbitrary point ($0 < M < \infty$), and III has to be fulfilled for some number ϵ (> 0), which may be chosen arbitrary small.

The question now arises if condition III is actually stronger than I and II together, or if it is only a consequence of our technique.

We can show that for potentials which in the neighborhood of a singularity can be given by a term Ax^{-k}, III is fulfilled if k is such that I and II are fulfilled.

Proof: (a) I is stronger than III at $x = 0$.
From I we have $|V(x)| < 1/x^k$ with $k < 1$. III gives $\int_0^M x^\epsilon x^{-(1-\alpha)(1+\epsilon)}\, dx < \infty$, with $k = 1 - \alpha$. We have $k' = 1 - \alpha\epsilon - \alpha$ and this is smaller than 1 for $\epsilon < 1$.

(b) II is stronger than III at infinity.
From II we have $|V(x)| < 1/x^k$ with $k > 1$. III gives: $\int_M^\infty x^\epsilon\, x^{(1+\alpha)(1+\epsilon)}\, dx < \infty$, $1 - \alpha = k$. This is finite if $1 + \alpha\epsilon + \alpha > 1$, or $0 < \epsilon < 1$.

(c) I and II together are sufficient for III at any intermediate point x_0, ($x_0 \neq 0$, $x_0 \neq \infty$). From I or II: $\int_{x_0-a}^{x_0+b} |V(x)|\, dx < \infty$.

From III: $\int_{x_0-a}^{x_0+b} |V(x)|^{1+\epsilon}\, dx < \infty$, if we assume that $V(x)$ has an insulated singularity at x_0.

A transformation ($y = x_0 - x$) brings the singularity at the origin, where we are left with the sum of two integrals of the form

$$\int_0^a |V(x)|^{1+\epsilon}\, dx.$$

But from I we have $|V(x)| \leq 1/x^k$, where $k \leq l < 1$. In order that also III exists, we must have $k + \epsilon \leq l + \epsilon < 1$; this can always be done by choosing a proper ϵ. Q.E.D.

For an integral equation of the type (5) in a normed vector space, the Fredholm alternative applies, i the kernel is completely continuous. That is, either

(i) the resolvent $R = (1 - KV)^{-1}$, $||R|| < \infty$ exists,

or

(ii) the corresponding homogeneous equation (system)

$$\psi = KV\psi \tag{7}$$

has a nontrivial solution.

Correspondingly, the domain of the k values, which is, so far, only the real axis, is divided into two complementary sets:

(i) the domain of the resolvent, which contains those k values for which the first alternative is valid, and

(ii) the spectrum which consists in the k eigenvalues of the operator KV.

We show below, that we do not leave C if we consider the homogeneous system (7) also for complex values of k in Im $k \geq 0$. In Appendix II we prove that our kernel $KV(k)$ is completely continuous in Im $k \geq 0$. Then the Fredholm alternative provides a way to prove uniqueness of the solutions of (5). One has only to show that the spectrum of (7) does not contain any real points.

C. Bound States (Eigenvalues)

The aim of this subsection is naturally to show that there are no real k eigenvalues except for $k = 0$.

Let us, for any k in Im $k \geq 0$, consider the solutions of the homogeneous equation (7).

Theorem 1. All $\psi \in C$ which allow an estimate of the kind (α = Im k)

$$|\psi(\mathbf{x})| \leq A(e^{-\alpha x}/x) \quad \text{for} \quad x \to \infty$$

form together a linear subspace $T \subset C$, which is mapped into itself under KV.

Proof: Let $\varphi = KV\psi$, $\psi \in T$. Then $|\varphi(\mathbf{x})| \leq ||KV||.\ |\psi(\mathbf{x})| \leq A(e^{-\alpha x}/x)$, since $||KV||$ is bounded.

Theorem 2. All solutions of (7) are in T.

To prove this we split V into a part V_0 which has finite range and a rest $\tilde{V}$.

$$V = V_0 + \tilde{V}. \tag{8}$$

With this we get

$$\psi = KV\psi = KV_0\psi + K\tilde{V}\psi.$$

We can choose V_0 so that $||K\tilde{V}|| < 1$ and it is also clear that $KV_0\psi = \psi_0 \in T$.

Then $\psi = \psi_0 + K\tilde{V}\psi$. But this is again an integral equation of the Fredholm type. Since $||K\tilde{V}|| < 1$, it can be solved by iteration, therefore, it has a unique solution in C as well as in T. But since $T \subset C$ they coincide; therefore $\psi \in T$.

In the case Im $k = \alpha > 0$ the eigenfunctions are square integrable and we can interpret them as bound states. Then with the well-known procedure one can also show that the k eigenvalues are pure imaginary. Carter[3] has shown that they lay between $0 \leq k \leq im$ and can only have a limiting point at $k = 0$. The possibility of a decomposition of V, as done in (8), is sufficient to prove that there are only finite-many singularities in Im $k \geq 0$.[4,5]

For the real part of the spectre we would like to show that it is empty except for the point $k = 0$. This would secure the unique solvability of the scattering equation.

The proof for a potential with finite range is easy and exactly the same as in the Schrödinger case.[1] The proof for an arbitrary central potential obeying conditions close to ours [I, II, III] has been given by Carter.[3] He uses the fact that, for a central potential, the Dirac equation can be separated, and then discusses the set of ordinary differential equations. Since this proof is very involved, we do not give it here.

For all the following, we make the assumption that the eigenvalues of the homogeneous scattering equation in the E plane lie on the real axis between $-m$ and m.

III. REGULARITY PROPERTIES OF THE T MATRIX

For a central potential which satisfies the conditions I, II, and III, the inhomogeneous scattering equation (5) has, for every real $k \neq 0$, exactly one solution which is continuous and bounded and composed by a plane-wave and a scattered-wave part;

[3] D. S. Carter, Thesis, Princeton University, Princeton, New Jersey, 1952 (microfilm).
[4] A. Grossman and A. T. Wu, J. Math. Phys. 2, 712 (1961).
[5] F. Riesz and B. S. Nagy, Vorl. Über Funktional Analysis, Nos. 66–76.

$$\psi = \phi + \psi_s. \tag{9}$$

The scattered-wave part behaves at least under the further condition[6]

$$\int_0^\infty y^2 \, |V(y)| \, dy < \infty \qquad \text{IV.}$$

for large x asymptotically as an outgoing spherical wave,

$$\psi_s \to r^{-1} \exp [ikr] f_{op} u \quad \text{for large} \quad r,$$

where

$$f_{op}(E, \vartheta) u(\mathbf{k}) = \frac{-1}{4\pi} (E + \beta m + i\boldsymbol{\alpha}\cdot\mathbf{k}') \times \int e^{-i\mathbf{k}\mathbf{y}} V(y) \psi(y) \, d^3y,$$

$$\mathbf{k}' = k(\mathbf{x}/x) \qquad \vartheta = \sphericalangle(\mathbf{k}', \mathbf{k}), \qquad E = \pm(k^2 + m^2)^{\frac{1}{2}}.$$

The T matrix, also an operator in spinor space, is derived from the transition matrix:

$$M_{fi} = (\phi_f, V\psi_i) = u_f T u_i.$$

Using (9), this gives the following representation:

$$T_{fi} = V(\mathbf{k}' - \mathbf{k}) + (\phi_f, V\psi_s). \tag{10}$$

Here $V(\mathbf{k}' - \mathbf{k})$ denotes the first Born approximation and is just the Fourier transform of the potential with respect to the momentum transfer between initial and final plane wave.

We wish to discuss the analytic behavior of $T(E, \cos\theta)$ as a function of the energy and the cosine of the scattering angle, or of other suitable scattering parameters. To this end we first investigate the behavior of $\psi(\mathbf{k}, \mathbf{x})$ as a function of (complex) k.

A. Regularity Properties of the Resolvent

Under the above conditions we may write

$$\psi(\mathbf{k}) = R(k)\phi_0(\mathbf{k}), \tag{11}$$

where R satisfies the Lippmann–Schwinger equation: $R = 1 + KVR$. The three functions in (11) are elements of a normed vector space and not just complex numbers. But it is well known[1,5] that the concept of analyticity can be applied to elements of a Banach space, practically by just replacing the absolute value sign by the norm sign. Most theorems of ordinary analysis, and all which we need for the following, can just be transcribed in that way.

[6] Our condition IV is necessary at this point since the Lemma of Pringsheim cannot be applied. [See Pringsheim, Math. Ann. 68, 367 (1910)].

Lemma 1. Regularity of $KV(k)$.

The following inequality holds for any k, k' if their imaginary part is bigger than or equal to zero:

$$|e^{ikr} - e^{-ik'r}| \leq |k - k'| \, r.$$

Therefore,

$$||KV(k) - KV(k')|| \leq |k - k'|^2 \int_0^\infty y^2 \, |V(y)| \, dy + |k - k'| \int_0^\infty y \, |V(y)| \, dy.$$

These integrals are bounded under conditions I and IV. Consequently, $KV(k)$ is a continuous function of k in Im $k \geq 0$. Carrying out a contour integration in this region, and applying Fubini's theorem, which is allowed since the integrals converge uniformly in Im $k > 0$, one shows that $KV(k)$ is regular in Im $k > 0$.

Lemma. 2. Regularity of $R(k)$.

If k_0 is not an eigenvalue of the scattering equation, then $R(k_0)$ exists, and because of the continuity of $KV(k)$, we have a complete neighborhood N of k_0 in which

$$|KV(k) - KV(k_0)| \leq 1/2 \, |R(k_0)| \quad \text{for all} \quad k \in N.$$

Let

$$\begin{aligned} S &= R(k_0)[1 - KV(k)] \\ &= 1 - R(k_0)[KV(k) - KV(k_0)] \\ &= 1 - H. \end{aligned}$$

For $k \in N$, $||H|| < \frac{1}{2}$ and S^{-1} can therefore be given by a Neumann series

$$S^{-1} = 1 + H + H^2 + \cdots .$$

Since this series converges absolutely and uniformly for all $k \in N$,

$$R(k) = S^{-1}(k) R(k_0)$$

is regular in the whole interior of N. But we already know where $R(k)$ exists, therefore, we obtain the following theorem:

$R(k)$ is regular in the half-plane Im $k > 0$, *except for finite-many singularities on the imaginary axis, which correspond to the bound states.*

B. Regularity of the Wavefunction ψ

For complex k, exp (ikx) is no longer bounded and the space C is too narrow for the solutions of the inhomogeneous scattering equation. We therefore consider the space C' containing all continuous

functions $\psi'(\mathbf{x})$ for which the following norm is finite:

$$||\psi'|| = \sup_{\mathbf{x},i} |e^{-\alpha x}\psi'(\mathbf{x})|.$$

α is a fixed positive number which we fix later. The correspondence

$$\psi(\mathbf{x}) = e^{-\alpha x}\psi'(\mathbf{x}) \tag{12}$$

introduces an isometric mapping from C on C' and vice versa. Now we understand the scattering equation as a functional equation in C' and construct, with (12), the corresponding equation in C:

$$\psi(\mathbf{x}) = e^{i\mathbf{k}\cdot\mathbf{x}-\alpha x} - \frac{e^{-\alpha x}}{4\pi}\int d^3y\,\frac{e^{ik|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x}-\mathbf{y}|}\,V(y)e^{\alpha y}\psi(\mathbf{y}). \tag{13}$$

Everything we had up to now is valid for this new equation, if only we make $V'(\mathbf{x}) = V(x)e^{\alpha x}$ subject to the same conditions as earlier $V(y)$; therefore

$$\int_0^\infty y^2 e^{\alpha y}\,|V(y)|\,dy < \infty. \qquad \text{V.}$$

This is our condition on $V(y)$ and α.

The resolvent of (13) is in C and has the already known regularity and continuity properties. The same is true for the resolvent $R'(k)$ of the scattering equation in C' because of (12):

$$\phi'(\mathbf{k}) = u'(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}}.$$

U' itself, as spinor amplitude of the free-particle solution of the Dirac equation, is a regular function of E. It has, however, branch points as a function of k at $k = \pm im$. With these exceptions, $\phi'(k)$ is a regular function of k with values in C' in the region $|\mathrm{Im}\ k| < \alpha$. Therefore,

$$\psi'(\mathbf{k}) = R(k)\phi'(\mathbf{k})$$

is also regular in

$$|\mathrm{Im}\ k| < \alpha, \qquad \mathrm{Im}\ k > 0, \tag{14}$$
$$k \neq \text{eigenvalue}.$$

Using (9), we can give the following results:

$\psi'_s(\mathbf{k}, \mathbf{x})$ is, for fixed $\mathbf{x}$, regular in $\mathbf{k}$ in the region (14), and for any fixed $\mathbf{k}$ in this region as a function of $\mathbf{x}$, continuous and bounded (in C').

C. Discussion of the T Matrix

We introduce as a new parameter, half of the momentum transfer ($\boldsymbol{\Delta}$).

$$\boldsymbol{\Delta} = \tfrac{1}{2}(\mathbf{k}' - \mathbf{k}) = \Delta\mathbf{e}', \qquad \mathbf{e}\cdot\mathbf{e}' = 0,$$
$$\mathbf{P} = \tfrac{1}{2}(\mathbf{k}' + \mathbf{k}) = P\mathbf{e}, \qquad P^2 = E^2 - \Delta^2 - m^2.$$

Here $\mathbf{e}$ and $\mathbf{e}'$ are unit vectors. The kernel $KV(k)$ and the T operator are at first only given for real values of E and Δ and for

$$E > +(m^2 + \Delta^2)^{\frac{1}{2}}.$$

Now we want to consider both functions also for complex values of the variables E and Δ. Since they are originally defined as functions of $k = \pm(E^2 - m^2)^{\frac{1}{2}}$, we have to make cuts into the complex-energy plane in order to define them unambiguously. The cuts run from $-\infty \to -m$ and from $+m \to +\infty$, and we choose the Riemannian sheet, so that the imaginary part of k is nonnegative ($\mathrm{Im}\ k \geq 0$). Just above the real axis, the real part of k is positive for $E > m$ and negative for $E < -m$. On the real axis for $-m \leq E \leq m$, k is pure imaginary.

Corresponding to (10), we separate the T matrix into the first Born approximation and a rest:

$$T = T_0(\Delta) + T_1(E, \Delta),$$
$$T_0(\Delta) = V(\mathbf{k}' - \mathbf{k}) = \int d^3y\, e^{-2i\boldsymbol{\Delta}\cdot\mathbf{y}}V(y).$$

This is an even function only of Δ and regular on the strip;

$$|\mathrm{Im}\ \Delta| < \tfrac{1}{2}\alpha.$$

To discuss $T_1(E, \Delta)$, we make a transformation with the whole scattering equation depending on a real parameter λ. This allows us the most profit possible from the exponential decrease of Green's function,

$$\psi^\lambda(\mathbf{x}) \equiv \psi(\mathbf{x})e^{i\lambda k\mathbf{e}\cdot\mathbf{x}}.$$

The transformed scattering equation becomes

$$\psi^\lambda = \phi^\lambda + K^\lambda V\psi^\lambda,$$
$$\phi^\lambda(\mathbf{x}) = \exp\,[-i\Delta\mathbf{e}\cdot\mathbf{x} + i(P + \lambda k)\mathbf{e}\cdot\mathbf{x}].$$

K^λ differs from K only in the exponent;

$$\exp\,(K^\lambda) = ik[|\mathbf{x} - \mathbf{y}| + \lambda\mathbf{e}\cdot(\mathbf{x} - \mathbf{y})].$$

We can transfer all the old results on the new scattering equation if the real part of the exponent of $K^\lambda(x, y)$ is not positive, i.e. if

$$-1 \leq \lambda \leq +1.$$

So $\phi^\lambda(E, \boldsymbol{\Delta}, \mathbf{x})$ is an element of C' if

$$[\mathrm{Im}\ \Delta)^2 + [\mathrm{Im}\ (P + \lambda k)]^2 < \alpha^2. \tag{15}$$

Then it follows, in the same way as (14),

$\psi^\lambda_s(E, \boldsymbol{\Delta}, \mathbf{x})$ is, for every fixed x, an analytic function in (15), except for two branch points at $E =$

$+(m^2 + \Delta^2)^{\frac{1}{2}}$—*the two cuts along the real E axis and the finite-many singularities in E between $-m$ and m which correspond to the bound states.*

For every pair (E, Δ) in this region, $\psi_s^\lambda(E, \Delta, \mathbf{x})$ as a function of x is continuous and bounded.

For $T_1(E, \Delta)$ we get the following representation:

$$T_1(E, \Delta) = (\phi_f, V\psi_s^\lambda) = (\phi_f, VK^\lambda V\psi_\lambda)$$

$$T_1(E, \Delta) = \frac{-1}{4\pi}(E + \beta m + i\boldsymbol{\alpha}\cdot\mathbf{k}') \times \int d^3x \exp[-i\Delta\mathbf{e}'\cdot\mathbf{x} - i(P + \lambda k)\mathbf{e}\cdot\mathbf{x}] \times V(x)\psi_s^\lambda(E, \Delta, \mathbf{x}).$$

From the preceding discussion, it is also clear that (15) gives just the condition that this integral exists.

Up to the often-mentioned singularities, which come from the resolvent $R(E)$, $T_1(E, \Delta)$ is regular in the following regions:

$$(\text{Im } \Delta)^2 + \{\text{Im}[(E^2 - \Delta^2 - m^2)^{\frac{1}{2}} + \lambda(E^2 - m^2)^{\frac{1}{2}}]\}^2 < \alpha^2, \qquad (16)$$

and

$$-1 \leq \lambda \leq +1.$$

We see that it is necessary that $|\text{Im } \Delta| < \alpha$. Then all pairs (E, Δ) belong to such a region for which, furthermore,

$$|\text{Im } (E^2 - \Delta^2 - m^2)^{\frac{1}{2}}| \leq |\text{Im } (E^2 - m^2)^{\frac{1}{2}}|,$$

because then λ can be chosen such that the second term in (16) vanishes. If, on the other side,

$$|\text{Im } (E^2 - \Delta^2 - m^2)^{\frac{1}{2}}| > |\text{Im } (E^2 - m^2)^{\frac{1}{2}}|,$$

then this term takes on a minimum for $\lambda = \pm 1$ and this minimum is

$$[|\text{Im } (E^2 - \Delta^2 - m^2)^{\frac{1}{2}}| - |\text{Im } (E^2 - m^2)^{\frac{1}{2}}|]^2.$$

With this we can give our result:

$T_1(E, \Delta)$ is regular in E and Δ in the region $|\text{Im } \Delta| < \alpha$, and

$$|\text{Im } |(E^2 - \Delta^2 - m^2)^{\frac{1}{2}}| - |\text{Im } (E^2 - m^2)^{\frac{1}{2}}| \leq (\alpha^2 - (\text{Im } \Delta)^2)^{\frac{1}{2}}$$

up to the two cuts and the finite-many singularities between $-m$ and m. (17)

As an application of the above result, let us show the form of the domains of analyticity in two special cases.

1. Regularity in $\cos\theta$ for fixed $E > m$

For real $E > m$, the region of regularity (17) of $T_1(E, \Delta)$ is characterized by

$$(\text{Im } \Delta)^2 + [\text{Im } (E^2 - \Delta^2 - m^2)^{\frac{1}{2}}]^2 < \alpha^2.$$

The relativistic connection between scattering angle and momentum transfer is

$$\Delta^2 = \tfrac{1}{2}(E^2 - m^2)(1 - \cos\theta).$$

Introducing this we find, for the domain (17),

$$[\text{Im } (\tfrac{1}{2}(1 - \cos\theta))^{\frac{1}{2}}]^2 + [\text{Im } (\tfrac{1}{2}(1 + \cos\theta))^{\frac{1}{2}}]^2 < \alpha^2/(E^2 - m^2),$$

$$(\text{Im } \sin\tfrac{1}{2}\theta)^2 + (\text{Im } \cos\tfrac{1}{2}\theta)^2 < \alpha^2/(E^2 - m^2).$$

Let us put $\theta = \varphi + i\psi$; then $\cos\theta = \cosh\psi\cos\varphi - i\sinh\psi\sin\varphi$, and this gives only a condition on ψ:

$$\sinh^2\tfrac{1}{2}\psi < \alpha^2/(E^2 - m^2).$$

Similar to the Schrödinger case for fixed ψ, $\cos\theta$, as a function of φ, describes an ellipse with focal points at ± 1. The boundary ellipse which contains the region of regularity is given by $\sinh\frac{1}{2}\psi = \alpha^2/(E^2 - m^2)$.

2. Regularity in Momentum Transfer

For which Δ may E vary in the whole, bounded E plane without violating (17)? (17) has to be valid for $E = 0$ and that means $|\Delta| < \alpha$.

We wish to show that this is also sufficient. Put $\Delta = r\exp(i\varphi)$. Then for all E,

$$|\text{Im } (E^2 - \Delta^2 - m^2)^{\frac{1}{2}}| - |\text{Im } (E^2 - m^2)^{\frac{1}{2}}| < [r^2 - (\text{Im } \Delta)^2]^{\frac{1}{2}}.$$

Proof: Introduce a complex parameter $Z = a + ib$:

$$(E^2 - m^2 - \Delta^2)^{\frac{1}{2}} = i\Delta\cosh Z,$$

$$(E^2 - m^2)^{\frac{1}{2}} = i\Delta\sinh Z,$$

$$\text{Im } (E^2 - m^2 - \Delta^2)^{\frac{1}{2}} = r(\cosh a\cos b\cos\varphi - \sinh a\sin b\sin\varphi),$$

$$\text{Im } (E^2 - m^2)^{\frac{1}{2}} \quad r(\sinh a\cos b\cos\varphi - \cosh a\sin b\sin\varphi).$$

Now we have to verify that, for all a and b, the following inequality is true:

$$|\cosh a\cos b\cos\varphi - \sinh a\sin b\sin\varphi| \leq |\cos\varphi| + |\sinh a\cos b\cos\varphi - \cosh a\sin b\sin\varphi|.$$

After squaring both sides,

$$\cos^2 b\cos^2\varphi - \sin^2 b\sin^2\varphi \leq \cos^2\varphi + 2|\cdots|.$$

This inequality evidently holds.

T_1 (E, Δ) is, therefore, regular in E and Δ in the direct product of the circle $|\Delta| < \alpha$ with the finite E plane, except the singularities (17).

IV. REMARKS

Reviewing, we see that our conditions on the potential are naturally more restrictive than in the Schrödinger case, but they are weaker than the conditions in reference 2.

There is one remark of caution to be made. Since our discussion proves only regularity in the finite E plane, we cannot say anything about the behavior of an expression like

$$\lim_{|E|\to\infty} T(E).$$

This is, however, necessary for the proof of dispersion relations. Khuri[2] has shown that

$$\lim_{|E|\to\infty} [T(E)/E] = 0,$$

at least if the Born approximation converges. Unlike the Schrödinger case, it is not proven up to now that, for high energies, the Born series converges, and to what expression it goes in the relativistic case.

ACKNOWLEDGMENTS

I would like to thank Professor R. Jost for the suggestion of the problem, Dr. W. Hunziker for continued guidance and advice, as well as Professor J. Sucher for many helpful discussions.

APPENDIX I. THE NORM OF KV

Let us put $K = K_1 + K_2$, where

$$K_1 = \left(E + \beta m + k\alpha_i \frac{x_i - y_i}{|\mathbf{x} - \mathbf{y}|}\right) G_0,$$

$$K_2 = -i\alpha_i \frac{x_i - y_i}{|\mathbf{x} - \mathbf{y}|} \cdot \frac{1}{4\pi} \frac{G_0}{|\mathbf{x} - \mathbf{y}|}.$$

A. Norm of the Operator K_1V

Note that

$$\left|\frac{x_i - y_i}{|\mathbf{x} - \mathbf{y}|}\right| \leq 1 \quad \text{and} \quad \left|\left|\alpha_i \frac{x_i - y_i}{|\mathbf{x} - \mathbf{y}|}\right|\right|$$

is smaller than some constant N. With this we see that

$$\left|\left|\int K_1 V\psi \, d^3y\right|\right| \leq \int |K_1 V\psi| \, d^3y <$$

$$||\psi|| \, |(E + \beta m + Nk)| \int |G_0 V| \, d^3y.$$

Since we only consider central potentials, we can always choose the coordinate system so that $V(\mathbf{x}) = V(x)$ and $|V(x)| = F(x)$.

With this we have

$$\int |G_0 V| \, d^3y \leq \int_0^\infty dy \, y^2 F(y) \frac{1}{4\pi} \int \frac{d\Omega}{|\mathbf{x} - \mathbf{y}|}.$$

Applying the identity

$$\frac{1}{4\pi} \int d\Omega \, f(|\mathbf{x} - \mathbf{y}|) = \frac{1}{2xy} \int_{|x-y|}^{x+y} rf(r) \, dr, \qquad \text{I.}$$

we get

$$\frac{1}{4\pi} \int \frac{d\Omega}{|\mathbf{x} - \mathbf{y}|} = \min\left(\frac{1}{x}, \frac{1}{y}\right) \leq \frac{1}{y},$$

and

$$\int |G_0 V| \, d^3y \leq \int_0^\infty dy \, yF(y).$$

$||K_1 V||$ is, therefore bounded if we demand that the potential $V(x)$ be a measurable function for which

$$|V(x)| \leq F(x) \quad \text{and} \quad \int_0^\infty xF(x) \, dx < \infty. \qquad \text{I.}$$

B. Norm of the Operator K_2V

Here too, the absolute value of the first factor is bounded by a constant N, therefore,

$$||K_2|| \leq \frac{N}{4\pi} \left|\left| \frac{\exp[ik(|\mathbf{x} - \mathbf{y}|)]}{|\mathbf{x} - \mathbf{y}|^2} \right|\right|.$$

Let us put $\varphi = K_2 V\psi$. Then $||\varphi(x)|| \leq ||\psi|| \cdot ||H(\mathbf{x})||$, where

$$H(\mathbf{x}) = \frac{N}{4\pi} \int \frac{|V(y)|}{|\mathbf{x} - \mathbf{y}|^2} \, d^3y,$$

or

$$H(\mathbf{x}) = N \int_0^\infty dy \, y^2 \, |V(y)| \cdot \frac{1}{4\pi} \int \frac{d\Omega}{|\mathbf{x} - \mathbf{y}|^2}.$$

Again $|V(x)| \leq F(x)$; then with the identity (I.1) we get

$$H(\mathbf{x}) \leq \frac{N}{2} \int_0^\infty dy \, F(y) \frac{y}{x} \log \frac{x + y}{|x - y|},$$

or

$$H(\mathbf{x}) \leq \frac{N}{2} \int_0^\infty F(y) \left|1 - \frac{y}{x}\right| \log \frac{1 + y/x}{|1 - y/x|} + \frac{N}{2} \int_0^\infty dy \, F(y) \log \frac{1 + y/x}{|1 - y/x|}.$$

But, with $S = y/x$, we have for all $S > 0$,

$$|1 - S| \log |(1 + S)/(1 - S)| < N_1 \text{ const.}$$

Therefore,

$$H(\mathbf{x}) < N_2 \int_0^\infty dy\, F(y) + N_3 \int_0^\infty dy\, F(y) \log \frac{1 + y/x}{|1 - y/x|},$$

where we have absorbed all the uninteresting constants into N_2 and N_3. We give an estimate of the second integral, using Hoelder's inequality:

$$\int_0^\infty dy\, F(y) \log \frac{1 + y/x}{|1 - y/x|} = \int_0^\infty (y^n F(y)) \left(y^{-n} \log \frac{1 + y/x}{|1 - y/x|} \right) dy$$

$$\leq \left[\int_0^\infty dy\, y^{np} F^p(y) \right]^{1/p} \times \left[\int_0^\infty dy\, y^{-nq} \left(\log \frac{1 + y/x}{|1 - y/x|} \right)^q \right]^{1/q},$$

where p and q are subject to the two conditions

$$1 < p, q < \infty,$$

and

$$1/p + 1/q = 1.$$

First we choose $q > 1$, but arbitrary otherwise. Then take $n = q^{-1}$, and substitute in the second integral $S = y/x$. The integral then reads

$$\mathrm{I} = \left[\int_0^\infty \frac{dS}{S} \left(\log \frac{1 + S}{|1 - S|} \right)^q \right]^{1/q}.$$

But this is bounded for any $q > 0$, because

(a) for $S \to 0$, the integral goes like S^{q-1}.

(b) for $S \to \infty$, the integral goes like $S^{-(1+q)}$.

(c) for $S \to 1$, the integral goes like $(\log x)^q$, for $x \to 0$.

In the first integral, we put $P = 1 + \epsilon$ and, therefore, $P/q = \epsilon$, where $\epsilon > 0$, but may be chosen arbitrary small.

This shows, that $||K_2 V||$ is bounded if we have the additional conditions on $V(x)$ $[|V(x)| \leq F(x)]$

$$\int_0^\infty F(x)\, dx < \infty, \qquad \text{II.}$$

and

$$\int_0^\infty y^\epsilon F(y)^{1+\epsilon}\, dy < \infty, \qquad \epsilon > 0. \qquad \text{III.}$$

The norm of the operator KV is, therefore, bounded if the potential $V(x)$ is measurable and fulfills the conditions I, II and III.

APPENDIX II. PROOF THAT $KV(k)$ IS COMPLETELY CONTINUOUS

To prove that KV is completely continuous, we first approximate it in norm with the aid of special potentials:

$$V_n(x) = \begin{cases} V(x) \text{ for } x \leq R_n \text{ and } |V(x)| \leq M_n, \\ 0 \quad \text{elsewhere.} \end{cases} \qquad \text{(II.1)}$$

We discuss R_n and M_n later. $KV = K_1 V + K_2 V$ as before;

$$||K_1 V_n - K_1 V|| \leq N_1 \int_{E_n} xF(x)\, dx,$$

and

$$||K_2 V_n - K_2 V|| \leq N_2 \int_{E_n} dy\, F(y) + N_3 \int_{E_n} dy\, y^\epsilon F(y)^{1+\epsilon},$$

where

$$E_n = \{y \mid y > R_n, \quad \text{or} \quad F(y) > M_n\}.$$

One can choose here R_n and M_n, one after the other, such that the right sides of the above equations become smaller than $1/n$. Therefore,

$$\lim_{n \to \infty} ||KV_n - KV|| = 0.$$

If KV is completely continuous, it maps every bounded set into a compact set E:

$$E = \{\varphi \mid \varphi = KV\psi, ||\psi|| \leq 1\}. \qquad \text{(II.2)}$$

We prove (II.2) for the special potentials (II.1). To show that E is compact, it is enough to show that the functions $\varphi \in E$ are all

(a) uniformly bounded,

(b) uniformly continuous at any point $\mathbf{a}$,

and, because our region is infinite,

(c) for large x are uniformly majorized by a function $g(x)$, for which

$$\lim_{x \to \infty} g(x) = 0,$$

(a) is evident from our proof of the boundedness of the operator KV. (b) was already shown for $K_1 V$ in reference 1. There it was proven that the first derivatives are uniformly bounded. This cannot be done in the same way with $K_2 V$, since it contains a term $|\mathbf{x} - \mathbf{y}|^{-2}$. If one differentiates φ under the integral sign with respect to $\mathbf{x}$, the integral no longer exists. We therefore approximate first $K_2(r)$ by a continuous differentiable function $K_3(r)$.

Take, for instance, ($r = |\mathbf{x} - \mathbf{y}|$):

$$K_\delta(r) = \frac{1}{4\pi} e^{ikr} f_\delta(r),$$

$$f_\delta(r) = \begin{cases} r^{-2} & \text{for } r \geq \delta, \\ \delta^{-2}(2 - r^2\delta^{-2}) & \text{for } r \leq \delta; \end{cases}$$

then

$$\lim_{\delta \to 0} ||K_2 V - K_\delta V|| = 0.$$

Proof:

$$||K_2 V - K_\delta V|| \leq \sup_{\mathbf{x}} \frac{M}{4\pi} \int_{r<\delta} r^{-2}\, d^3y + \sup_{\mathbf{x}} \frac{M}{4\pi} \int_{r<\delta} \delta^{-2}(2 - r^2\delta^{-2})\, d^3y,$$

since the contribution for $r \geq \delta$ cancel. Solving the integrals,

$$||K_2 V - K_\delta V|| \leq M\delta + \frac{2}{\delta^2} \frac{M}{4\pi} \tfrac{4}{3}\pi\delta^3 = \tfrac{5}{3} M\delta.$$

Since $\varphi = KV\psi$, we have $||\varphi_2 - \varphi_{2\delta}|| \leq \frac{5}{3}M\delta$, where $\varphi_{2\delta} = K_\delta V\psi$, and $\varphi_2 = K_2 V\psi$. The functions $\varphi_{2\delta}(r)$ are continuously differentiable and their derivative can be formed under the integral sign. We get

$$|\partial K_\delta/\partial r| \leq C_\delta < \infty, \quad \text{for all } r, \quad \text{and}$$

therefore,

$$|\partial \varphi_\delta(\mathbf{x})/\partial x_k| < \tfrac{1}{3} MR^3 C_\delta \quad \text{for all } \mathbf{x}.$$

With this we can show the uniform continuity. Given $\epsilon > 0$, we first choose $\delta > 0$ so that $\frac{5}{3}M\delta < \frac{1}{3}\epsilon$. With this, C_δ is fixed. Then we can choose a number $\eta > 0$ so that $\frac{1}{3}MR^3 C_\delta \eta < \frac{1}{3}\epsilon$. Now if only $|\mathbf{x} - \mathbf{a}| < \eta$, we have

$$|\varphi(\mathbf{x}) - \varphi(\mathbf{a})| \leq |\varphi(\mathbf{x}) - \varphi_\delta(\mathbf{x})| + |\varphi_\delta(\mathbf{x}) - \varphi_\delta(\mathbf{a})| + |\varphi_\delta(\mathbf{a}) - \varphi(\mathbf{a})| \leq \tfrac{1}{3}\epsilon + \tfrac{1}{3}\epsilon + \tfrac{1}{3}\epsilon = \epsilon,$$

for all $\mathbf{a}$ and all $\varphi \in E$.

Therefore, the functions $\varphi \in E$ are uniformly continuous, even for all $\mathbf{a}$, which is more than we required.

(c) For all $\varphi \in E$: $||\varphi|| = ||KV\psi|| \leq ||KV||$, and this is

$$\leq M\,|(E + \beta m + Nk)|\, \frac{1}{4\pi} \int_{K_R} \frac{d^3y}{r} + \frac{NM}{4\pi} \int_{K_R} \frac{d^3y}{r^2}.$$

We are only interested in the region $x > R$; there,

$$(|\mathbf{x} - \mathbf{y}|)^{-1} \leq (x - y)^{-1} \leq x^{-1}(1 - x^{-1}y)^{-1} < x^{-1}C_1$$

analog

$$(|\mathbf{x} - \mathbf{y}|)^{-2} \leq C_2 x^{-2};$$

therefore,

$$||\varphi|| \leq M\,|(E + \beta m + Nk)|\,(C_1/x)\tfrac{1}{3}R^3 + MN(C_2/x^2)\tfrac{1}{3}R^3 \equiv g(x).$$

Q.E.D.

The Nature of the Axioms of Relativistic Quantum Field Theory. II*

E. C. G. SUDARSHAN
Department of Physics and Astronomy, University of Rochester, Rochester, New York
(Received 25 February 1963)

This paper continues the study of the nature and interdependence of the axioms of relativistic field theory; attention is focused on the notion of relativistic invariance. The central result of the present paper is the derivation of necessary and sufficient conditions for a representation of the covariant free field to admit a unitary representation of the inhomogeneous Lorentz group associated with the field operator. It is shown that only the standard Fock–Cook representation has this property. The relevance of the requirement that the Lorentz group is represented by a unitary family associated with the field operator is exhibited by an analysis of the covariant representations of Shale and Segal. These representations involve extremal states which are not pure, and group representations by intertwining operators.

I. INTRODUCTION

IN relativistic quantum field theory[1] it is natural to assume, in accordance with the principle of relativity, that the correspondence $\phi(x) \rightarrow \phi(\Lambda x + a)$ of the field operators, imaging the change of frame represented by (a, Λ), must leave the algebra of the field operators unaltered, i.e., it must be an automorphism. These automorphisms of the operator algebra preserve Hermiticity properties and constitute a realization of the (inhomogeneous) Lorentz group. In conventional treatments of relativistic quantum field theory, it is also assumed that there must exist a family of unitary operators $U(a, \Lambda)$ furnishing a (unitary, true) representation of the (inhomogeneous) Lorentz group and implementing the local automorphism:

$$\phi(x) \rightarrow U(a, \Lambda)\phi(x)U^{-1}(a, \Lambda) = \phi(\Lambda x + a).$$

Since the representations are assumed to be continuous, one can assert the existence of the ten Hermitian infinitesimal generators of $U(a, \Lambda)$, and impose additional spectral conditions on these generators. It is usually assumed that there exists a unique invariant state called the vacuum, and that all other states belong to the continuous positive energy spectrum of the Hamiltonian (time-translation) operator. Of course there is no *a priori* reason to insist that the unitary family $U(a, \Lambda)$ must implement the local automorphism. We may, then, distinguish several distinct concepts that enter the characterization of a relativistic field:

(i) The local automorphism $\phi(x) \rightarrow \phi(\Lambda x + a)$.

(ii) A unitary family $U(a, \Lambda)$ furnishing representation of the Lorentz group.

(iii) The local automorphism $\phi(x) \rightarrow \phi(\Lambda x + a)$ being implemented by the unitary family $U(a, \Lambda)$.

(iv) The existence of the (unique) vacuum.

(v) The nonnegative spectrum of the Hamiltonian.

While it is usual to include the requirements of a unique vacuum and of a nonnegative energy spectrum as well as of the local automorphism implemented by a unitary family under the postulate of relativistic invariance, they are by no means essential. We know of models in which the vacuum is not unique,[1,2] and nontrivial interacting models of quantum field theory exist in which the unitary family $U(a, \Lambda)$ fulfills the representation and spectrum conditions, but does not lead to the local automorphism.

It is also generally assumed that the field operator ring contains the unitary family $U(a, \Lambda)$; loosely speaking this implies that the ten generators can be "built up" using the field operators. This apparently innocent axiom has the consequence that all theories with an invariant cyclic vacuum are in fact direct integrals of theories with a unique invariant vacuum.[3] In the major part of the investigations in the sequel we shall assume that the family $U(a, \Lambda)$ is contained in the operator ring generated by the

* Supported in part by the U. S. Atomic Energy Commission.

[1] For part I, see E. C. G. Sudarshan and K. Bardakci, J. Math. Phys. **2**, 767 (1961).

[2] It is amusing to note at this point that not only is the uniqueness of the vacuum not essential, but the vacuum itself may be dispensed with. The simplest example is provided by starting with the conventional free neutral scalar field $\varphi(x)$ in the standard Fock–Cook representation and constructing the Wightman polynomial $\phi(x) = :\varphi^2(x):$ which can be shown to be a local field. [A. S. Wightman, *Cours de la Faculte des Sciences de l'Université de Paris*, 1957–8; p. 57 (unpublished)]. It may be seen to be irreducible over all odd (even) particle states of the free field $\varphi(x)$. But in the set of all states of $\varphi(x)$ with an odd number of quanta, there is no invariant state! In spite of this, $\phi(x)$ is a local field defined over these states and undergoing local automorphisms implemented by a unitary family.

[3] T. F. Jordan and E. C. G. Sudarshan, J. Math. Phys. **3**, 587 (1962); D. Ruelle, Helv. Phys. Acta **35**, 147 (1962); H. J. Borchers, "On Structure of the Algebra of Field Operators" (Institute for Advanced Study, Princeton, New Jersey, preprint).

field operator, but discuss a class of field theories in which this is not so.

The important question to be investigated is whether every relativistic field theory admitting the local automorphism has a unitary family $U(a, \Lambda)$ implementing it. There has recently been considerable interest in the nonstandard representations in field theory and in the possibility of breakdown of formal symmetry properties of the theory in the actual realization. It is then natural to ask if the unitary family $U(a, \Lambda)$ exists for the nonstandard representations of any theory. In the following sections, we investigate this question in detail for a free neutral scalar field: the result is somewhat unexpected. Within the framework of an *irreducible* representation of the operator algebra, only the standard Fock–Cook representation admits such a unitary family.[4] This result, obtained within the axiomatic framework outlined, is most simply stated as follows: "Of all representations of the relativistic (neutral, scalar) free field, only the standard representation obtains a manifestly covariant local unitary transformation." The role of the irreducibilty assumption is seen from the results of Shale and Segal[5] discussed in detail in Sec. 5; that, if it is relaxed, there exists a one-parameter infinity of theories in which a unitary family $U(a, \Lambda)$ implementing the local automorphism exist, but the spectrum conditions are violated in that the Hamiltonian is not positive definite (except for the Fock–Cook representation).

II. AUTOMORPHISMS AND REPRESENTATIONS OF THE FIELD RING

Let $\mathfrak{L}$ be the (one-particle) Hilbert space of square integrable functions $\psi(\mathbf{r})$ of the three-vector variable $\mathbf{r}$. Then a unitary representation of the (proper, orthochronous) inhomogeneous Lorentz group is furnished by the choice[6]

$$\begin{aligned}(h\psi)(\mathbf{r}) &= \int \omega(\mathbf{r}, \mathbf{r}')\psi(\mathbf{r}')\, d^3r' \\ (\mathbf{p}\psi)(\mathbf{r}) &= -i\nabla\psi(\mathbf{r}) \\ (\mathbf{j}\psi)(\mathbf{r}) &= -i(\mathbf{r}\times\nabla)\psi(\mathbf{r}) \\ (\mathbf{k}\psi)(\mathbf{r}) &= \frac{1}{2}\int (\mathbf{r}+\mathbf{r}')\omega(\mathbf{r}, \mathbf{r}')\psi(\mathbf{r}')\, d^3r'.\end{aligned} \tag{2.1}$$

for the ten generators h, $\mathbf{p}$, $\mathbf{j}$, $\mathbf{k}$ of the group. We have used the kernel $\omega(\mathbf{r}, \mathbf{r}')$ defined by

$$\omega(\mathbf{r}, \mathbf{r}') = (2\pi)^{-3}\int \exp\,[i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')](\mathbf{q}^2+m^2)^{\frac{1}{2}}\, d^3q.$$

This representation is irreducible[7] and corresponds to a particle of spin zero and mass m. We shall denote the finite unitary transformation corresponding to a Lorentz transformation (a, Λ) obtained from these ten generators by the unitary operator $R(a, \Lambda)$ in the space $\mathfrak{L}$.

Let $u_\alpha(\mathbf{r})$ be an orthonormal basis in $\mathfrak{L}$ so that

$$\int u_\alpha^*(\mathbf{r})u_\beta(\mathbf{r})\, d^3r = \delta_{\alpha\beta}. \tag{2.2}$$

We define the kernel $\mathcal{C}(\mathbf{r}, \mathbf{r}')$ by

$$\mathcal{C}(\mathbf{r}, \mathbf{r}') = 2^{-\frac{1}{2}}(2\pi)^{-3} \times \int \exp\,[i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')](\mathbf{q}^2+m^2)^{-\frac{1}{4}}\, d^3q.$$

Let a_α be a sequence of (unbounded) operators and let a_α^+ be their adjoints which satisfy the commutation relations

$$[a_\alpha, a_\beta^+] = \delta_{\alpha\beta}; \qquad [a_\alpha, a_\beta] = 0. \tag{2.3}$$

We shall call the ring generated by the operators the field ring. Then the (relativistic, neutral scalar) field operator is given by the construction[8]

$$\phi(\mathbf{r}, t) = \sum_\alpha \int \mathcal{C}(\mathbf{r}, \mathbf{r}') \exp\,(iht) \times [a_\alpha u_\alpha(\mathbf{r}') + a_\alpha^+ u_\alpha^*(\mathbf{r}')]\, d^3r'. \tag{2.4}$$

The field operator $\phi(x) = \phi(\mathbf{x}, x_0)$ then satisfies the commutation relation

$$\begin{aligned}[\phi(x), \phi(x')] &= i\Delta(x-x') \\ &= -i(2\pi)^{-3}\int (\mathbf{q}^2+m^2)^{\frac{1}{2}} \exp\,(i\mathbf{k}\cdot\mathbf{y}) \times \sin\,\{y_0(\mathbf{q}^2+m^2)^{\frac{1}{2}}\}\, d^3q.\end{aligned}$$

Under the transformation

$$u_\alpha(\mathbf{r}) \to R(a, \Lambda)u_\alpha(\mathbf{r}),$$

the field operator $\phi(x)$ transforms locally:

$$\phi(x) \to \phi(\Lambda x + a).$$

[4] This statement refers only to the free field. The essential point is not the existence of the vacuum (compare reference 2), but the existence of the energy operator; see Sec. 3.

[5] I. E. Segal, Illinois J. Math. **6**, 500 (1962).

[6] See, for example, L. L. Foldy, Phys. Rev. **102**, 568 (1956).

[7] This representation is not equivalent to the reducible representation obtained from a (local) relativistic wave equation with a (manifestly) covariant amplitude. To see this explicitly, we note that a local relativistic wave equation is invariant not only under the real Lorentz transformations, but also under complex Lorentz transformations. It is, in particular, invariant under the antichronous proper transformation $\mathbf{r} \to -\mathbf{r}$, $t \to -t$. But under this operation, frequencies change sign; hence the "energy" must also change sign for the one-particle amplitude. This is of course true in the familiar spin-0, spin-1/2 and spin-1 covariant wave equations.

[8] For a detailed discussion see, for example, S. S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Row, Peterson and Company, Evanston, Illinois, 1961) Sec. 7.

By virtue of the invariance of the commutation relations under the change $x \to \Lambda x + a$, $x' \to \Lambda x' + a$, it follows that the local transformation $\phi(x) \to \phi(\Lambda x + a)$ is an automorphism.

We may now consider this automorphism of the field operator as being generated by a linear automorphism of the field ring [rather than by a linear transformation on the $u_\alpha(\mathbf{r})$]. For this purpose let us write

$$R_{\alpha\beta}(a, \Lambda) = \int u_\alpha(\mathbf{r}) R(a, \Lambda) u^*_\beta(\mathbf{r})\, d^3r. \tag{2.5}$$

Then the local transformation of the field operators is equivalent to the linear automorphism

$$a_\alpha \to \sum_\beta R_{\alpha\beta} a_\beta; \qquad a^+_\alpha \to \sum_\beta R^*_{\alpha\beta} a^+_\beta \tag{2.6}$$

of the field ring. The question of the representation of the field operator is the same as that of the representation of the field ring.

The interesting question to be discussed now is whether these automorphisms of the field ring can be generated as inner automorphisms, i.e., whether there exists an operator family $U(a, \Lambda)$ such that

$$U(a, \Lambda) a_\alpha U^{-1}(a, \Lambda) = \sum_\beta R_{\alpha\beta}(a, \Lambda) a_\beta. \tag{2.8}$$

Since Hermiticity relations are preserved by these automorphisms, if the representation of the field ring is irreducible, the $U(a, \Lambda)$ would be a unitary family (apart from an unessential scalar). Even in the more general reduction into factors,[9] if $U(a, \Lambda)$ belongs to the field ring, the same property holds.

III. LORENTZ COVARIANCE OF IRREDUCIBLE REPRESENTATIONS OF THE FIELD RING

In this section we wish to investigate the automorphism of the field operators more closely. We shall be particularly interested in the necessary and sufficient conditions under which there may exist an inner automorphism of the operator algebra for every Lorentz transformation (a, Λ). Let e be any one (or linear combination) of the ten generators h, $\mathbf{p}$, $\mathbf{j}$, $\mathbf{k}$. Define

$$e_{\alpha\beta} = \int u^*_\alpha(\mathbf{r}) e u_\beta(\mathbf{r})\, d^3r \tag{3.1}$$

in terms of the basic set of one-particle wavefunctions $u_\alpha(\mathbf{r})$. Let us also assume that we have an irreducible representation (or, more generally, a factor representation) of the field ring.

We shall now explicitly assume that $U(a, \Lambda)$ belong to this ring.[9] Let $U(\tau)$ be any one-parameter family belonging to $U(a, \Lambda)$. Since $U(\tau)$ constitute a one-parameter unitary family on the Hilbert space $\mathfrak{H}$ on which the field operators are represented, by Stone's theorem,[10] there exists a Hermitian generator E for this family in $\mathfrak{H}$ which satisfies the relations

$$U(\tau) = 1 + i\tau E + O(\tau^2),$$
$$U(\tau) a_\beta U^{-1}(\tau) = a_\beta + i\tau [E, a_\beta] + O(\tau^2).$$

If $U(\tau)$ is the unitary family corresponding to the generator e in the one-particle Hilbert space $\mathfrak{L}$, it follows then that

$$a_\alpha \to \sum_\beta R_{\alpha\beta}(\tau) a_\beta = a_\alpha + i\tau \sum_\beta e_{\alpha\beta} a_\beta + O(\tau^2).$$

We have, on comparing the two transformations to first order in τ,

$$[E, a_\alpha] = \sum_\beta e_{\alpha\beta} a_\beta. \tag{3.2}$$

Consider the operator

$$F = \sum_{\alpha,\beta} e_{\alpha\beta} a^+_\alpha a_\beta; \qquad [F, a_\alpha] = \sum_\beta e_{\alpha\beta} a_\beta,$$

if it exists. Then it follows that (if F exists),

$$[E - F, a_\alpha] = 0,$$

so that in every irreducible (or factor) realization of the oscillator ring, $E - F$ must be a scalar, which we denote by C. Hence if the operator F exists, the generator E has the form

$$E = \sum_{\alpha,\beta} e_{\alpha\beta} (a^+_\alpha a_\beta + \mathfrak{C}_{\alpha\beta}), \tag{3.3}$$

where $\mathfrak{C}_{\alpha\beta}$ may be chosen arbitrarily except insofar as to require

$$C = \sum_{\alpha,\beta} e_{\alpha\beta} \mathfrak{C}_{\alpha\beta}.$$

On the other hand, we may show that if this structure (3.3) for E does not exist, there can be no operator associated with the oscillator ring which satisfies (3.2). To see this, let us consider the index α to be restricted to take on a finite set of values. Then the corresponding condition (3.2) states that

$$[E, a_\alpha] = \sum_\beta e_{\alpha\beta} a_\beta; \qquad \alpha \in \{\alpha_1, \cdots, \alpha_N\}.$$

The right-hand side exists by definition of the automorphism. This is satisfied only if

$$E = \sum_\alpha a^+_\alpha (\sum_\beta e_{\alpha\beta} a_\beta) + \mathcal{E}, \tag{3.4}$$

[9] Compare T. F. Jordan and E. C. G. Sudarshan, reference 3.

[10] M. H. Stone, Ann. Math. **33**, 643 (1932); J. von Neumann, Ann. Math. **33**, 567 (1932); F. Riesz and B. Sz-Nagy, "Functional Analysis," translated from French by L. F. Boron, (Frederick Ungar Publishing Company, New York, 1955), p. 383.

where

$$[\mathcal{E}, a_\alpha] = 0,$$

i.e., $\mathcal{E}$ is associated with the oscillator ring a_γ, a_γ^+ where γ does not assume any of the N values $\alpha_1, \cdots, \alpha_N$. Since the subset $\alpha_1, \cdots, \alpha_N$ is arbitrary, this can be true if and only if E has the form (3.4). Hence we have proved that the necessary and sufficient condition for the existence of the operator E associated with the field ring is that, for a suitable set of constants $\mathcal{C}_{\alpha\beta}$, there exists a nontrivial operator E of the form (3.3).

Let now E, E' be the operators corresponding to the one-particle operators e, e' belonging to some group of continuous automorphisms. This implies that E, E' have the structures

$$E = \sum_{\alpha,\beta} e_{\alpha\beta}(a_\alpha^+ a_\beta + \mathcal{C}_{\alpha\beta}),$$

$$E' = \sum_{\alpha,\beta} e'_{\alpha\beta}(a_\alpha^+ a_\beta + \mathcal{C}'_{\alpha\beta}),$$

with suitable constants $\mathcal{C}_{\alpha\beta}$, $\mathcal{C}'_{\alpha\beta}$. Then we have

$$[E, E'] = \sum_{\alpha,\beta}\sum_{\alpha',\beta'} e_{\alpha\beta}e'_{\alpha'\beta'}[a_\alpha^+ a_\beta, a_{\alpha'}^+ a_{\beta'}]$$

$$= \sum_{\alpha,\beta} [e, e']_{\alpha\beta} a_\alpha^+ a_\beta. \tag{3.5}$$

It is important to note that on the right-hand side of (3.5) there are no constant terms. Hence the correspondence between e and E expressed by (3.5) may now be further restricted by stating that if $e = [e', e'']$, then $\mathcal{C}_{\alpha\beta} = 0$ in the expression (3.5) for E. Hence for such E we have

$$E = \sum_{\alpha,\beta} e_{\alpha\beta} a_\alpha^+ a_\beta. \tag{3.6}$$

For the case of the Lorentz group, the ten generators h, $\mathbf{p}$, $\mathbf{j}$, $\mathbf{k}$ have this property; it follows that for any of these operators, the corresponding generator associated with the oscillator ring is given by expressions of the form (3.6). Hence the necessary and sufficient conditions for the Lorentz invariance of the theory, i.e., for the existence of a unitary family $U(a, \Lambda)$ associated with the oscillator ring, is that the ten quantities H, $\mathbf{P}$, $\mathbf{J}$, $\mathbf{K}$ defined by the equations

$$H = \sum_{\alpha,\beta} h_{\alpha\beta} a_\alpha^+ a_\beta,$$

$$\mathbf{P} = \sum_{\alpha,\beta} \mathbf{p}_{\alpha\beta} a_\alpha^+ a_\beta,$$

$$\mathbf{J} = \sum_{\alpha,\beta} \mathbf{j}_{\alpha\beta} a_\alpha^+ a_\beta,$$

$$\mathbf{K} = \sum_{\alpha,\beta} \mathbf{k}_{\alpha\beta} a_\alpha^+ a_\beta, \tag{3.7}$$

all exist.

We now observe that only in the standard representation of the (finite-mass m, spin-0) field can the Hamiltonian exist, since

$$H \geq \sum_{\alpha,\beta} m\delta_{\alpha\beta} a_\alpha^+ a_\beta = m \sum a_\alpha^+ a_\alpha; \tag{3.8}$$

but the right-hand side is infinite (i.e. does not exist) for any representation except the standard one.[11] It thus follows that, in the case of the relativistic (finite-mass, spin-0) field, none except the standard (Fock) representation is Lorentz-covariant. With unessential technical modifications, the proof can be adapted to any finite-mass free field.

IV. SHALE-SEGAL STATES AND REDUCIBLE COVARIANT REPRESENTATIONS

The Wightman formulation[12] of the free field is well known. It suffices here to say that in this formulation the existence of the ten generators of the Lorentz group and the existence of an invariant state (vacuum) are postulated. The analysis in the previous sections asserts that the nonstandard representations of the free field do not fall within the Wightman framework.

The relevant point here is that, by virtue of a familiar construction,[12,13] if we can define a linear functional over the field ring which is left invariant under the automorphism of the oscillator ring, the theory furnishes a unitary representation of the group of automorphisms. We outline the proof of this assertion.

Let Ω be a linear functional over the oscillator ring which is invariant under the Lorentz automorphisms $A \to \Omega(A)$, and the collection of linear functionals Ω_B defined by $A \to \Omega_B(A) = \Omega(BA)$, for any operators A, B. Then we can define a Hilbert space with a standard state ω and operators $\mathcal{O}_B$ associated with the elements B of the algebra defined by

$$\Omega \to \omega; \qquad B \to \mathcal{O}_B; \qquad \Omega_B \to \omega_B,$$

with the representation

$$\mathcal{O}_B \omega_A = \omega_{BA}$$

[11] For readable accounts, see A. S. Wightman and S. S. Schweber, Phys. Rev. **98**, 812 (1955); R. Haag, *Lectures on Theoretical Physics* edited by W. E. Brittin, B. W. Downs and J. Downs (Interscience Publishers, Inc., New York, 1961), Vol. III.

[12] A. S. Wightman, Phys. Rev. **101**, 860 (1956).

[13] I. E. Segal, Bull. Am. Math. Soc. **53**, 73 (1947); I. Gelfand and M. Naimark, Mat. Sbornik **54**, 197 (1943) (in Russian); M. A. Naimark, *Normed Rings*, translated by L. F. Boron, (P. Noordhoff Ltd., Groningen, The Netherlands, 1959). See also R. Haag and B. Schroer, J. Math. Phys. **3**, 248 (1962).

for the operators $\mathcal{O}_B$. The scalar product is defined as

$$(\omega_B, \omega_A) = \Omega(B^+ A).$$

Let us now consider the automorphism of the oscillator ring associated with $R(a, \Lambda)$. Let $A \to A'$, etc. under this automorphism. Then $\omega_A \to \omega_{A'}$ etc.; but the scalar product becomes

$$(\omega_B, \omega_A) \to (\omega_{B'}, \omega_{A'}) = \Omega(B'^+ A') = \Omega(B^+ A) = (\omega_B, \omega_A),$$

so that there exists the true *unitary* family $U(a, \Lambda)$ in the Hilbert space which yields

$$U(a, \Lambda)\omega_A = \omega_{A'}.$$

It satisfies, in particular, the property of leaving the standard state invariant:

$$U(a, \Lambda)\omega = \omega.$$

The representation is a true representation, and by Stone's theorem[10] there exists generators for every one-parameter family. In particular, the ten generators of the Lorentz group all exist. However, *there is no assurance that $U(a, \Lambda)$ belongs to the oscillator ring.* Nor is it guaranteed that the oscillator ring has an irreducible (a factor) representation. But it is true that the "vacuum state" ω is cyclic with respect to the oscillator ring.

If $U(a, \Lambda)$ did belong to the oscillator ring, and if the vacuum is cyclic, then all irreducible (factor) representations into which the given representation may be decomposed have an invariant vacuum state, and the unitary family $U(a, \Lambda)$ simultaneously decomposes.[9] Then, by virtue of the results above, it would follow that all the irreducible (factor) representations must be the standard (Fock) representation of the oscillator ring.

Two remarkable results concerning linear functionals invariant under automorphisms have been presented by Segal[5]; we state the results here without proof (and suitably paraphrased):

Theorem (Shale): There exists an infinite one-parameter family of invariant linear functionals on the oscillator ring and associated inequivalent representations of the oscillator ring.

Theorem (Segal): Any universally invariant linear functional is a convex integral of these fundamental linear functionals. In every one of these inequivalent representations, except the standard Fock representation, the generator, associated with a one-parameter automorphism corresponding to a positive-definite one-particle generator, has a partially negative spectrum. In particular, the Hamiltonian is not positive definite.

In the Shale theorem, the universal invariance refers to an *arbitrary* linear automorphism corresponding to an arbitrary unitary transformation in the space of one-particle wavefunctions. If we restrict ourselves to the linear automorphisms corresponding to the Lorentz group, it may be necessary to weaken the theorem by omitting the second part of the theorem. Shale fundamental linear functionals are defined as follows: Let A be any operator associated with a finite subset of oscillator variables

$$\{a_{\beta_1}, \cdots, a_{\beta_N}; a^+_{\beta_1}, \cdots, a^+_{\beta_N}\}.$$

Let $D(n; \beta_1, \cdots, \beta_N)$ be the projection operator associated with the operator $\sum_{i=1}^{N} a^+_{\beta_i} a_{\beta_i}$ corresponding to the eigenvalue n. Then consider the linear functional that assigns the numerical value

$$E_C(A) = (1 - C)^N \sum_{n=0}^{\infty} C^n \operatorname{tr} \{A D(n; \beta_1, \cdots, \beta_N)\}, \tag{4.1}$$

where tr corresponds to the trace relative to the finite subset of oscillators and $0 \leq C < 1$. The universal invariance of the linear functional is apparent since the projection operator $D(n; \beta_1, \cdots, \beta_N)$, as well as the operation of relative trace, are invariant under arbitrary linear automorphisms of the finite set of oscillator variables, corresponding to a finite-dimensional unitary transformation.

Observing that the Shale linear functionals are invariant under the Lorentz automorphisms, making use of the linear functional construction, we have a unitary representation of the Lorentz group on a Hilbert space and an invariant state. If the unitary family $U(a, \Lambda)$ belonged to the oscillator ring, the known result on the reduction of representations of fields with an invariant vacuum state then assert that the representation is reducible,[3] and further, that in each of the reduced representations there exists an invariant state.[9] But since $E_C(1) = 1$,

$$E_C\Big(\sum_\alpha a^+_\alpha a_\alpha\Big) = \lim_{N\to\infty} (1 - C)^N \times \sum_{n=0}^{\infty} C^n \cdot n \cdot \operatorname{tr} \{D(n; \beta_1, \cdots, \beta_N)\}$$

$$= \lim_{N\to\infty} (1 - C)^N \sum_{n=0}^{\infty} \frac{n C^n (N + n - 1)!}{(N - 1)! n!} = \infty, \tag{4.2}$$

it follows that not all of them can be standard representations. *Hence the unitary family $U(a, \Lambda)$ does not belong to the oscillator ring.*

These results point out that the Shale states furnish a new class of representations of an operator

algebra by linear operators, and its automorphisms by unitary operators, in terms of direct integrals of representations of the algebra which do not in general, furnish a representation of the automorphisms. The unitary operators representing the automorphisms do not leave the component representations of the operator algebra invariant, but in fact intertwine these component representations. We are then led to conjecture that the Shale states are not pure states even though they are extremal elements of the convex set of universally invariant states. To verify this conjecture, let us restrict ourselves to a finite subset of oscillator variables; then the density matrix representing the state is

$$\rho_N = (1 - C)^N \sum_{n=0}^{\infty} C^n D(n; \beta_1, \cdots, \beta_N), \qquad (4.3)$$

so that $\text{tr}\,\{\rho_N\} = 1$, but

$$\begin{aligned}\text{tr}\,\{\rho_N^2\} &= (1 - C)^{2N} \sum_{n=0}^{\infty} C^{2n}\,\text{tr}\,\{D(n; \beta_1, \cdots, \beta_N)\} \\ &= (1 - C)^{2N}(1 - C^2)^{-N} \\ &= [(1 - C)/(1 + C)]^N \neq 1 = \text{tr}\,\{\rho_N\},\end{aligned}$$

for any value of N.

V. DISCUSSION

We thus find that the various aspects of relativistic invariance of quantized field theories imply different things, and to a large extent, these requirements are independent. We may have a local automorphism $\phi(x) \to \phi(\Lambda x + a)$ but no unitary operator $U(a, \Lambda)$ in a particular representation of the field operators; it then means that it is meaningless to talk about an energy–momentum operator and the spectral conditions. This comes about since in an irreducible representation of a set of operators, it is not automatic that groups of (linear) automorphisms of the operators get represented; in general such an automorphism generates an inequivalent representation of the operator algebra. Since automorphisms of a Hamiltonian dynamical system are called canonical transformations, we see that not all canonical transformations are unitary transformations. It is curious to observe that in the logical structure of dynamics, the primitive dynamical attributes of energy, momentum, and angular momentum are associated with automorphisms (canonical transformations) of the dynamical variables, rather than directly with functions of the dynamical variables themselves.

On the other hand, the existence of a unitary family $U(a, \Lambda)$ representing Lorentz transformations does not imply the *local* manifestly covariant transformation:

$$\phi(x) \to U(a, \Lambda)\phi(x)U^{-1}(a, \Lambda) \neq \phi(\Lambda x + a). \qquad (5.1)$$

Such a theory may be constructed as follows: Choose the standard (Fock) representation of the free field. Then we can explicitly construct the projection operator to a two-particle state, following a construction of von Neumann.[14] We have

$$\mathcal{P}(\alpha, \beta) = \tfrac{1}{2}a_\alpha^+ a_\beta^+ \mathcal{P}(O) a_\alpha a_\beta. \qquad (5.2)$$

Here $\mathcal{P}(O)$ is the vacuum state projection operator:

$$\begin{aligned}\mathcal{P}(O) = \prod_{\alpha=1}^{\infty} (2\pi)^{-1} &\int_{-\infty}^{\infty} dx_\alpha \int_{-\infty}^{\infty} dy_\alpha \\ &\times \exp\,[(x_\alpha + iy_\alpha)/\sqrt{2}a_\alpha] \\ &\times \exp\,[(x_\alpha - iy_\alpha)/\sqrt{2}a_\alpha^+], \qquad (5.3)\end{aligned}$$

which does not vanish by definition of the standard representation. Let $V(\alpha_1\ \beta_1;\ \alpha_2\beta_2)$ be the Möller matrix for a relativistic interacting two-particle system. Such unitary Möller matrices exist.[15] Now construct the field operator:

$$\begin{aligned}\psi(x) = \{1 - \sum_{\alpha_1,\beta_1} \mathcal{P}(\alpha_1, \beta_1)\}\phi(x)\{1 - \sum_{\alpha_2,\beta_2} \mathcal{P}(\alpha_2, \beta_2)\} \\ + \tfrac{1}{2} \sum_{\alpha_1,\beta_1} \sum_{\alpha_2,\beta_2} V^*(\alpha_1\beta_1; \alpha_2\beta_2)\mathcal{P}(\alpha_2, \beta_2)\phi(x) \\ + \tfrac{1}{2} \sum_{\alpha_1,\beta_1} \sum_{\alpha_2,\beta_2} V(\alpha_1\beta_1; \alpha_2\beta_2)\phi(x)\mathcal{P}(\alpha_2, \beta_2). \qquad (5.4)\end{aligned}$$

This field operator is unitarily equivalent to the field operator $\phi(x)$, since the transformation $\phi(x) \to \psi(x)$ is equivalent to the unitary transformation in the Hilbert space of the field operator in which the "two-particle" states undergo the unitary transformation by the Möller matrix $V(\alpha_1\beta_1; \alpha_2\beta_2)$. The resulting theory leads to nontrivial scattering in the two-particle channel, and only in that channel; it is hence highly artificial. On the other hand, the transformation of $\psi(x)$ when $\phi(x)$ transforms by $U(a, \Lambda)$, is by the family

$$\begin{aligned}U'(a, \Lambda) = U(a, \Lambda)\{1 - \sum_{\alpha,\beta} \mathcal{P}(\alpha, \beta)\} \\ + \frac{1}{4} \sum_{\alpha_1,\beta_1} \sum_{\alpha_2,\beta_2} \sum_{\alpha_1',\beta_1'} \sum_{\alpha_2',\beta_2'} V(\alpha_1'\beta_1'; \alpha_2'\beta_2') \\ \times \mathcal{P}(\alpha_2', \beta_2')U(a, \Lambda)\mathcal{P}(\alpha_1, \beta_1)V(\alpha_1\beta_1; \alpha_2\beta_2). \qquad (5.5)\end{aligned}$$

These transformations are nonlocal, but are never-

[14] J. von Neumann, Math. Ann. **104**, 570 (1931). See also A. S. Wightman and S. S. Schweber, reference 8.

[15] T. F. Jordan, A. Macfarlane, and E. C. G. Sudarshan, "A Hamiltonian Model of Lorentz Invariant Particle Interactions," (to be published).

theless unitary, and are obtained by a unitary transformation on the family $U(a, \Lambda)$. Consequently, the family $U'(a, \Lambda)$ is a unitary representation of the Lorentz transformations and is in accord with the spectral conditions. However, since the transformations under Lorentz transformations are *nonlocal*, the elegant *analyticity* properties of the Wightman functions do not obtain for the "interacting" field.

We have already seen that the existence of local unitary automorphisms of the field operator does not imply the existence of a vacuum state (invariant linear functional) even if the spectrum conditions are satisfied. On the other hand, from the Segal theorem, we see that a unitary family of local automorphisms does not imply the spectrum conditions even if a vacuum state (invariant linear functional) exists.

The lack of Lorentz covariance of the nonstandard representations of the free relativistic field implies that the so-called "thermodynamic limit" of the free field (in which the particle density is finite over all space) is not Lorentz-covariant.

Instead of the relativistic field and covariance under the Lorentz group, we could consider other dynamical systems and other groups of automorphisms. One familiar example of this type is a spin assembly with a ferromagnetic Hamiltonian, i.e., an infinite number of localized "spins" (constituting a spatial lattice and with mutual interactions favoring parallel alignment of spins). There then exist states of infinite spin ("ferromagnetic states") with the resultant spin of the ferromagnet oriented along an arbitrary axis in space ("along the direction of the trace magnetic field"). On the other hand, since the (interacting) Hamiltonian of the spin assembly is rotationally invariant, no direction is preferred over any other; it is usually stated that the ground state must be infinitely degenerate since every one of these states with the "infinite" spin has the same energy. In the light of the results stated before, it is clear that to refer to this phenomenon as "degeneracy of the ground state" is misleading since each one of these states of infinite spin corresponds to a different representation; degeneracy refers to states in the same irreducible representation of the dynamical system. It also follows that while there is a rotation automorphism of the spin algebra (which leaves the ferromagnet Hamiltonian unchanged), the "infinite spin" states do not furnish a unitary representation of these automorphisms. In other words the "ferromagnet" is not rotation-covariant, and its angular momentum is undefined. We can, however, construct the Shale–Segal representations for the spin assembly in which ferromagnetic states are included; but in these representations the spin assembly is not irreducibly represented. In actual physical situations, one does not consider infinite spin assemblies, and it is clear that only a countable number of inequivalent representations exist (corresponding to all values of total spins up to a maximum finite spin, and these with suitable multiplicities). However, the restriction of the inequivalent ferromagnetic states to a finite number of spins forms a convenient starting point for a perturbation theory which may be useful below the Curie temperature. It is perhaps important to note that *the existence of the various representations is purely kinematic* (i.e., depending only on the operator structure of the dynamical system), and not on its dynamics (Hamiltonian); the dynamics merely help "stabilize" the states and make them occur in physically interesting applications.

It is tempting to believe that the considerations outlined here apply to the structure of the representations of interacting fields. For the trivial nonfree system of theories involving Wightman polynomials[2] (since normal ordering is still defined!), these considerations certainly apply. But no result of this kind is known for any genuine interacting theories. (Nor does one know if there are genuine interacting field theories!). For the time being, the relevance of these considerations to interacting fields must remain a hope.

APPENDIX A. REPRESENTATIONS OF THE FIELD RING

The question of the representation (see reference 9) of the field operator $\phi(x)$ is the same as the representation of the ring a_α, a_α^+. The most familiar representation of the field ring is the "Fock" representation furnished by all sequences of nonnegative integers $\{n_i\}$, with $\sum_{i=1}^{\infty} n_i < \infty$ considered as basis vectors of a Hilbert space so that

$$(\{n_i^{(1)}\}, \{n_i^{(2)}\}) = \prod_{i=1}^{\infty} \delta(n_i^{(1)}, n_i^{(2)}).$$

The oscillator variables a_α, a_α^+ have the representation

$$a_\alpha\{n_i\} = n_\alpha^{\frac{1}{2}}\{n_i - \delta_{\alpha_i}\};$$

$$a_\alpha^+\{n_i\} = (n_\alpha + 1)^{\frac{1}{2}}\{n_i + \delta_{\alpha_i}\},$$

so that the n_i can be thought of as occupation numbers, and a_α, a_α^+ as annihilation and creation operators. We shall refer to this representation as

the standard representation (or the Fock representation).

Inequivalent representations of indescribable multiplicity exist. An uncountable number of such representations is obtained by the above construction, but by relaxing the requirement that $\sum_{i=1}^{\infty} n_i$ be finite. We note that if $\{n_i^{(1)}\}$ and $\{n_i^{(2)}\}$ belong to the same representation, then $\sum_{i=1}^{\infty} |n_i^{(1)} - n_i^{(2)}| < \infty$. We may define equivalence of two sequences of nonnegative integers by requiring that $\sum_{i=1}^{\infty} |n_i^{(1)} - n_i^{(2)}|$ be finite; this is then reflexive, symmetric, and transitive, and consequently defines uncountably many equivalence classes of sequences of nonnegative integers. Each sequence defines a representation of the oscillator ring, and there are uncountably many inequivalent representations. All these representations are called "discrete."

A multitude of inequivalent representations of the oscillator ring can be obtained as follows: Consider the transformation

$$a_\alpha \to b_\alpha = \sum_\beta \{V(\alpha, \beta)a_\beta + W(\alpha, \beta)a_\beta^+\},$$

with

$$\sum_\beta \{V(\alpha, \beta)V^*(\alpha', \beta) - W(\alpha, \beta)W^*(\alpha', \beta)\} = \delta_{(\alpha, \alpha')}.$$

Then b_α, b_α^+ are also oscillator variables:

$$[b_\alpha, b_\beta^+] = \delta_{\alpha\beta}; \qquad [b_\alpha, b_\beta] = 0.$$

We can now construct the (uncountably many, inequivalent) discrete representations with respect to the oscillator variables. For almost all transformations V, W, these representations are inequivalent among themselves and in relation to the discrete representations with respect to the primitive variables. The simplest class of such representations is obtained by considering an infinite set $\{\nu_k^r\}$ of infinite subsets of the indices j, so that r and k run over an infinite set of values, and taking

$$V(\nu_k^r, \nu_l^s) = \delta_{rs}\delta_{kl} \cosh \theta(r),$$
$$W(\nu_k^r, \nu_l^s) = \delta_{rs}\delta_{kl} \sinh \theta(r).$$

There are uncountably many such choices of parameters, and the discrete states with respect to these oscillator variables are inequivalent.

A third class of representations is discussed in connection with the results of Segal and Shale in Sec. 5.

APPENDIX B. REPRESENTATION OF GUAGE TRANSFORMATIONS

There are cases in which a one-parameter group of automorphisms admit a unitary representation for one of the nonstandard representations of the oscillator ring. Consider the case of a *charged* (scalar) field $\psi(x)$ and the one-parameter group of gauge transformations of the first kind:

$$\psi(x) \to e^{i\lambda}\psi(x).$$

This is an automorphism since it leaves the commutation relations

$$[\psi(x), \psi^*(x)] = i\Delta(x - x')$$

invariant. Since we have a charged field, $\psi(x)$ is no longer Hermitian, and the expansion in terms of one-particle function introduces two sets of oscillators a_α, a_α^+; b_α, b_α^+. The automorphism on the oscillator ring is

$$a_\alpha \to e^{i\lambda}a_\alpha, \qquad b_\alpha \to e^{-i\lambda}b_\alpha,$$
$$a_\alpha^+ \to e^{-i\lambda}a_\alpha^+, \qquad b_\alpha^+ \to e^{i\lambda}b_\alpha^+.$$

For α running over a finite index set, this automorphism is generated by the unitary operator

$$V(\lambda) = \exp\{i\lambda \sum_\alpha (b_\alpha^+ b_\alpha - a_\alpha^+ a_\alpha)\}.$$

If the (unbounded) hermitian operator

$$Q = \sum_{\alpha=1}^{\infty} (a_\alpha^+ a_\alpha - b_\alpha^+ b_\alpha)$$

exists, then the unitary operator

$$U(\lambda) = \exp(-i\lambda Q)$$

also exists, and generates the automorphism. In other words, we have a realization of the gauge transformations of the first kind provided Q defined by (3.10) exists. However this is not a *necessary* condition since it is sufficient if the operator

$$Q' = \sum_{\alpha=1}^{\infty} (a_\alpha^+ a_\alpha - b_\alpha^+ b_\alpha + \mathfrak{C}_\alpha)$$

exists for suitable choice of the constants $\mathfrak{C}_\alpha$. Consequently, every one of the *discrete* representations of the oscillator ring is gauge-covariant. The question of the gauge covariance of the continuous representations is more complicated; in general they are not gauge-covariant. It is not known whether there are any gauge-covariant continuous representations of the oscillator ring.

ACKNOWLEDGMENT

The author is pleased to thank Dr. T. F. Jordan and Dr. J. M. Cook for their criticism of the manuscript.

Generalized "Cross-Correlation" Field Quantities as Solutions of Field Equations*

PHILIP R. KARR
Space Technology Laboratories, Inc., Redondo Beach, California
(Received 28 January 1963)

It has been shown by Wolf that, if a field amplitude obeys the wave equation, then a derived field quantity (the cross correlation or mutual coherence function), also obeys the wave equation. This result helps clarify the subject of partially coherent light; for example, a Huygens' principle for the propagation of *intensity* is a consequence. A generalization is presented: For that class of linear partial differential equations (p.d.e.'s) in which at least one independent variable, say t, does not appear in the coefficients, one can construct from a solution $f(P, t)$, (P representing collectively the other independent variables) a generalized "cross-correlation" function $F(P, t; P_1, t_1, P_2, t_2 \cdots)$ satisfying the same p.d.e. as $f(P, t)$; F is an integral over s of $f(P, t+s)\, g(s, P_1, t_1, P_2, t_2 \cdots)$, where $(P_1, t_1, P_2, t_2 \cdots)$ are values of (P, t); for particular choices of the highly arbitrary g, F can be, for example, (a) the usual cross-correlation function, used by Wolf, (b) a higher-order correlation function, (c) the Hilbert transform of $f(P, t)$, and (d) derivatives of $f(P, t)$. For linear p.d.e.'s in which two or more independent variables have the above property, higher-dimensional or vector versions of F are obtainable. The existence of this (two-fold) generalization of Wolf's result suggests the possibility of other physical applications.

1. INTRODUCTION

IN recent years there has been considerable interest in the theory of partial coherence of light as evidenced by a large number of papers on the subject. One of the prominent contributions has been some work by Wolf[1] concerning the propagation of the coherence function or the correlation function in a wave field. In particular, Wolf discovered that in a field governed by a wave equation (such as that of an optical disturbance), wherein the amplitude propagates in space and time, the correlation function of this disturbance also will obey the wave equation and propagates in a similar manner. Wolf made use of this result in various problems dealing with partially coherent light; in particular, he showed that a kind of Huygens' principle for the propagation of intensity is a consequence. In the present paper it is pointed out that Wolf's result is a special case of a far more general result.

First, considering the wave equation, many quantities *other* than the correlation function, which may be derived from a given solution, are also solutions of the original equation. For example, higher-order correlation functions also have this property. Second, this property may be extended to many equations *other* than the wave equation. It is hoped that by presenting this generalization of the previous result we may stimulate other applications in addition to those already made.

* Presented at meeting of the American Physical Society at Seattle, Washington, August 1962.

[1] M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, Inc., New York, 1959), Chap. 10. E. Wolf, Proc. Roy. Soc., London, **A230**, 246 (1955).

2. DEVELOPMENT

The previous work[1] considers the wave equation

$$\nabla^2 f = (1/c^2)(\partial^2 f/\partial t^2), \tag{1}$$

and considers a solution $f(P, t)$ where P represents, collectively, the spatial variables x, y, z. Forming the cross-correlation function or coherence function

$$\Gamma(P_1, t_1, P_2, t_2) = \text{av}\, f(P_1, t_1 + t) f^*(P_2, t_2 + t), \tag{2}$$

it is found that this also obeys Eq. (1), where (P_1, t_1) or (P_2, t_2) are taken to be the independent variables; i.e., (2) obeys the pair of equations

$$\nabla_i^2 \Gamma = (1/c^2)(\partial^2 \Gamma/\partial t_i^2) \qquad (i = 1, 2), \tag{3}$$

where ∇_i^2 is the Laplacian operator with respect to the coordinates P_i. In Eq. (2), av represents average with respect to t; (*) represents complex conjugate. As mentioned above, this result turns out to be useful; for example, Wolf combines it with the Kirchhoff–Helmholtz integral formula to derive a kind of Huygens' principle for the propagation of intensity in a wave field.

Now consider that class of operator equations

$$Lf = 0, \tag{4}$$

where the operator L acts on f, a function of independent variables $(x, y, z, w \cdots)$; we require the following two properties: (1) if $f(x, y, z, w \cdots)$ is a solution, then we can change at least one of the independent variables, say z by a constant c and still achieve a solution, in other words, $f(x, y, z + c, w \cdots)$ is also a solution; (2) Equation (4) is linear. If Eq. (4) has the above two properties, which may

be briefly called (1) the shifting property, and (2) linearity, then we may achieve a large generalization of Wolf's result. These two conditions are met by the class of linear partial differential equations for which at least one of the independent variables does not appear in the coefficients. A large number of the equations of physics are of this type including, of course, the wave equation (1).

Assume that linearity holds and that the shifting property is true with respect to the independent variable denoted t (not necessarily time), and denote, collectively, the other independent variables by P; then it is seen that if $f(P, t)$ is a solution of Eq. (4) then there is a generalized solution

$$F(P, t, P_1, t_1, P_2, t_2 \cdots) = \int ds\, f(P, t + s) g(s; P_1, t_1, P_2, t_2 \cdots). \quad (5)$$

By interpreting the highly arbitrary g and the limits of integration in various ways, we obtain, formally, various solutions F. We say "formally" since, of course, some of the indicated operations may, under certain conditions, produce divergent or mathematically "nonexistent" functions. The following are some examples of (5):

(a) If

$$g(s) = (-d/ds)\,\delta(s), \quad (6)$$

where $\delta(s)$ is the Dirac delta function, we obtain as a solution the derivative

$$F = (\partial/\partial t) f(P, t). \quad (7)$$

(b) If g is interpreted as $-1/\pi s$, and if the integration is taken between $-\infty$ and ∞ in a principal-value sense, then we have

$$F(P, t) = \frac{-1}{\pi} \int_{-\infty}^{\infty} ds\, \frac{f(P, t + s)}{s} = \frac{1}{\pi} \int_{-\infty}^{\infty} dt'\, \frac{f(P, t')}{t - t'}, \quad (8)$$

i.e., $F(P, t)$ is now the Hilbert transform[2] of $f(P, t)$, with respect to the variable t; since the Hilbert transform is defined for real functions, we have here taken $f(P, t)$ as real.

(c) If we take g as

$$g = \text{const}\, f^*(P_1, t_1 + s), \quad (9)$$

we have

$$F(P, t, P_1, t_1) = \text{const} \int ds\, f(P, t + s) f^*(P_1, t_1 + s). \quad (10)$$

Comparing Eq. (10) with Eq. (2), we see that this is essentially the cross-correlation function between the solutions for two different points, where the integral is taken in the "av" sense; that is, as the limit of the integral, ("const" being the reciprocal of the range in s, the range being extended to infinity).

It is to be noted that in Eq. (2) Wolf takes $f(P, t)$ as complex, with imaginary part the Hilbert transform of the real part. In (b) above, we have shown that the Hilbert transform of a solution of Eq. (1) or of Eq. (4) is a solution, so this form of $f(P, t)$ is encompassed in our treatment.

(d) By taking

$$g = \text{const}\, f(P_1, t_1 + s) f(P_2, t_2 + s) f(P_3, t_3 + s) \cdots, \quad (11)$$

we obtain

$$F(P, t; P_1, t_1, P_2, t_2 \cdots) = \text{const} \int ds\, f(P, t+s) f(P_1, t_1 + s) f(P_2, t + s) \cdots. \quad (12)$$

Expression (12) may be regarded as a generalized higher-order correlation function, so Eq. (5) thus includes this sort of generalization. If the shifting property holds with respect to more than one of the independent variables, higher-dimensional or vector versions of the generalized solution F are available.

3. CONCLUDING REMARKS

For that class of linear-operator equations (4) obeying the "shifting property" (also referred to as invariance under the translation group[3]), the "generalized cross-correlation function" F, Eq. (5), may be formally constructed from a specific solution; F is then also a solution.

Wolf's remark that the cross-correlation function in a wave field obeys the wave equation (1) is equivalent to a special case of our result. Perhaps other "physical" consequences of this generalized result can be added to those already found[1] for the case of partially coherent light. There may be cases

[2] J. Dugundji, IRE Trans. Info. Theory 4, 53 (1958); E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals* (Claremont Press, Oxford, England, 1948), Chap. 5.

[3] E. Goursat, E. R. Hedrick, and O. Dunkel, *A Course in Mathematical Analysis* (Ginn and Company, New York, 1917) Vol. II, Part II, pp. 86–99.

in which boundary conditions can be conveniently set up in terms of some form of F, instead of a more conventional "amplitude." Furthermore, analogues of the Kirchhoff–Helmholtz integral formula may be available in other areas. It may be of some interest to note that, by the use of this formula, one can set up a generalized Huygens' principle for the propagation of any nth power of the amplitude in a light field (n an interger) or of the nth moment of the amplitude.

ACKNOWLEDGMENTS

During the progress of this work, the writer has had the benefit of useful comments by some of his colleagues, including, among others, D. E. Royal, R. Graves, and P. Stehle.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 8 AUGUST 1963

Random Matrix Diagonalization—Some Numerical Computations*

C. E. PORTER
Brookhaven National Laboratory, Upton, Long Island, New York
(Received 5 April 1963)

Numerical results of Monte-Carlo calculations of spacing and eigenvalue distributions for the invariant and independent Gaussian orthogonal ensemble of Hamiltonian matrices are presented. Many of the histograms should be useful for comparison with experimental data. A table of the first few moments of each distribution is given. For the spacing distributions, such moments are equivalent to spacing correlation coefficients, and hence these are also made available indirectly.

I. INTRODUCTION

THE use of random matrix models to explain the observed statistical fluctuations in the energy level spacings, energy-level widths, and expectation values of complex spectra is now a well-established point of view.[1-12] However, the difficulties of extracting analytical results from such models to compare to experimental data are considerable, and this task has so far been accomplished only in special cases and with a tremendous display of analytical virtuosity.[3-9] It is the purpose of this article to present a complete set of numerical results for one model to indicate what might in the future be obtained analytically and to provide much needed theoretical results to compare to experimental data. (We do not go into the data analysis in this paper.)

II. THEORETICAL MODEL

The model used in this paper is the invariant, independent Gaussian orthogonal ensemble in 10 dimensions.[1] Thus the real symmetric Hamiltonian matrix (10×10) is distributed according to the form

$$P(H) = C \exp(-\mathrm{Tr}\, H^2/4\sigma^2). \quad (1)$$

where C is a normalization constant and σ is the root-mean-square dispersion of the off-diagonal matrix elements. The parameter σ is a scale factor, and is related to the mean distance D between levels by the approximate connection (valid for large N)

$$D = \pi\sigma/N^{\frac{1}{2}}, \quad (2)$$

where in this paper we are concerned with $N = 10$.

The details of the numerical computations are

* Research supported by the United States Atomic Energy Commission.

[1] For a complete list of early references, see C. E. Porter and N. Rosenzweig, Ann. Acad. Sci. Fennicae **AVI**, No. 44 (1960).

[2] N. Rosenzweig and C. E. Porter, Phys. Rev. **120**, 1698 (1960); Phys. Rev. **123**, 853 (1961).

[3] M. L. Mehta, Nucl. Phys. **18**, 395 (1960); M. L. Mehta and M. Gaudin, Nucl. Phys. **18**, 420 (1960); **22**, 340 (1960); M. Gaudin, Nucl. Phys. **25**, 447 (1960).

[4] F. J. Dyson, J. Math. Phys. **3**, 140 (1962); **3**, 157 (1962); **3**, 166 (1962); **3**, 1191 (1962); **3**, 1199 (1962).

[5] F. J. Dyson and M. L. Mehta, J. Math. Phys. **4**, 701 (1963); M. L. Mehta and F. J. Dyson, J. Math. Phys. **4**, 713 (1963).

[6] J. Gunson, J. Math. Phys. **3**, 752 (1962).

[7] K. G. Wilson, J. Math. Phys. **3**, 1040 (1962).

[8] L. D. Favro, P. B. Kahn, and M. L. Mehta, "Concerning Polynomials Encountered in the Study of the Distribution Function of Spacings," Brookhaven National Laboratory Report BNL 757 (T-280), September 1962 (unpublished).

[9] P. B. Kahn, Nucl. Phys. **41**, 159 (1963).

[10] N. Rosenzweig, Bull. Am. Phys. Soc. **7**, 91 (1962) and *Brandeis Summer Institute Lecture Notes* (W. A. Benjamin, Inc., New York, 1962).

[11] H. S. Leff, Bull. Am. Phys. Soc. **8**, 31 (1963).

[12] C. E. Porter, Phys. Letters **2**, 292 (1962); Nucl. Phys. **40**, 167 (1963); Physics Today **16**, No. 2, 26 (1963), **16** No. 4, 90 (1963).

in which boundary conditions can be conveniently set up in terms of some form of F, instead of a more conventional "amplitude." Furthermore, analogues of the Kirchhoff–Helmholtz integral formula may be available in other areas. It may be of some interest to note that, by the use of this formula, one can set up a generalized Huygens' principle for the propagation of any nth power of the amplitude in a light field (n an interger) or of the nth moment of the amplitude.

ACKNOWLEDGMENTS

During the progress of this work, the writer has had the benefit of useful comments by some of his colleagues, including, among others, D. E. Royal, R. Graves, and P. Stehle.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 8 AUGUST 1963

Random Matrix Diagonalization—Some Numerical Computations*

C. E. PORTER
Brookhaven National Laboratory, Upton, Long Island, New York
(Received 5 April 1963)

Numerical results of Monte-Carlo calculations of spacing and eigenvalue distributions for the invariant and independent Gaussian orthogonal ensemble of Hamiltonian matrices are presented. Many of the histograms should be useful for comparison with experimental data. A table of the first few moments of each distribution is given. For the spacing distributions, such moments are equivalent to spacing correlation coefficients, and hence these are also made available indirectly.

I. INTRODUCTION

THE use of random matrix models to explain the observed statistical fluctuations in the energy level spacings, energy-level widths, and expectation values of complex spectra is now a well-established point of view.[1-12] However, the difficulties of extracting analytical results from such models to compare to experimental data are considerable, and this task has so far been accomplished only in special cases and with a tremendous display of analytical virtuosity.[3-9] It is the purpose of this article to present a complete set of numerical results for one model to indicate what might in the future be obtained analytically and to provide much needed theoretical results to compare to experimental data. (We do not go into the data analysis in this paper.)

II. THEORETICAL MODEL

The model used in this paper is the invariant, independent Gaussian orthogonal ensemble in 10 dimensions.[1] Thus the real symmetric Hamiltonian matrix (10×10) is distributed according to the form

$$P(H) = C \exp(-\mathrm{Tr}\, H^2/4\sigma^2). \quad (1)$$

where C is a normalization constant and σ is the root-mean-square dispersion of the off-diagonal matrix elements. The parameter σ is a scale factor, and is related to the mean distance D between levels by the approximate connection (valid for large N)

$$D = \pi\sigma/N^{\frac{1}{2}}, \quad (2)$$

where in this paper we are concerned with $N = 10$.

The details of the numerical computations are

* Research supported by the United States Atomic Energy Commission.

[1] For a complete list of early references, see C. E. Porter and N. Rosenzweig, Ann. Acad. Sci. Fennicae **AVI**, No. 44 (1960).

[2] N. Rosenzweig and C. E. Porter, Phys. Rev. **120**, 1698 (1960); Phys. Rev. **123**, 853 (1961).

[3] M. L. Mehta, Nucl. Phys. **18**, 395 (1960); M. L. Mehta and M. Gaudin, Nucl. Phys. **18**, 420 (1960); **22**, 340 (1960); M. Gaudin, Nucl. Phys. **25**, 447 (1960).

[4] F. J. Dyson, J. Math. Phys. **3**, 140 (1962); **3**, 157 (1962); **3**, 166 (1962); **3**, 1191 (1962); **3**, 1199 (1962).

[5] F. J. Dyson and M. L. Mehta, J. Math. Phys. **4**, 701 (1963); M. L. Mehta and F. J. Dyson, J. Math. Phys. **4**, 713 (1963).

[6] J. Gunson, J. Math. Phys. **3**, 752 (1962).

[7] K. G. Wilson, J. Math. Phys. **3**, 1040 (1962).

[8] L. D. Favro, P. B. Kahn, and M. L. Mehta, "Concerning Polynomials Encountered in the Study of the Distribution Function of Spacings," Brookhaven National Laboratory Report BNL 757 (T-280), September 1962 (unpublished).

[9] P. B. Kahn, Nucl. Phys. **41**, 159 (1963).

[10] N. Rosenzweig, Bull. Am. Phys. Soc. **7**, 91 (1962) and *Brandeis Summer Institute Lecture Notes* (W. A. Benjamin, Inc., New York, 1962).

[11] H. S. Leff, Bull. Am. Phys. Soc. **8**, 31 (1963).

[12] C. E. Porter, Phys. Letters **2**, 292 (1962); Nucl. Phys. **40**, 167 (1963); Physics Today **16**, No. 2, 26 (1963), **16** No. 4, 90 (1963).

described elsewhere.[13] It is sufficient to note that a straightforward Monte-Carlo calculation was performed in which 10 000 random 10 × 10 matrices were generated according to the distribution (1) and then diagonalized, and the resulting spectra were sorted to yield the results which are plotted in the graphs and tabulated in Table I of this paper. The eigenvector components were, of course, obtained also, but since the eigenvector component distribution is well understood in this (invariant) case (for example, the marginal distribution of a single component is that of a component of a randomly oriented unit vector in 10 dimensions—see reference 1), we do not discuss the eigenvector components further since they feed naturally into the distributions of widths (transition probabilities) and expectation values which are discussed elsewhere.[2,14]

It has been shown by Mehta[5] that an intimate connection exists between the circular orthogonal and circular symplectic ensembles. The connection is such that the spacing distributions in the circular symplectic ensembles are identical with alternate spacing distributions in the circular orthogonal ensemble, i.e., the nearest-neighbor spacing distribution in the circular symplectic ensemble is the next-nearest-neighbor spacing distribution in the circular orthogonal ensemble with the mean nearest-neighbor spacing of the circular orthogonal ensemble doubled, etc.

It is suspected that a similar connection exists, at least in the infinite-dimensional limit, between the Gaussian symplectic and the Gaussian orthogonal ensembles, but this has not yet been proven. If this connection does exist, then our computations for the Gaussian orthogonal ensemble may apply to the Gaussian symplectic ensemble as well.

TABLE I. Tabulation of Moments.

	Moment	1	2	3	4
Spacings	Neighbor				
	0	1.001	1.316	2.105	3.925
	1	2.001	4.484	11.09	29.93
	2	3.001	9.590	32.48	116.1
	3	4.001	16.67	72.21	324.4
	4	5.001	25.73	136.1	738.7
	5	6.001	36.78	230.1	1468.
	6	7.001	49.84	360.6	2651.
	7	8.001	64.90	533.7	4448.
	8	9.001	82.02	756.7	7065.
Eigenvalues	All	−.001347	.4327	−.0004436	.2515
	1, 10	.8404	.7257	.6432	.5845
	2, 9	.6047	.3795	.2463	.1648
	3, 8	.4156	.1845	.08655	.04257
	4, 7	.2455	.07075	.02264	.007850
	5, 6	.1066	.01701	.003376	.0007729

III. NUMERICAL RESULTS

The numerical results are shown in Figs. 1–6. In Figs. 1–3 are plotted the histograms for the nine possible spacing distributions P^k for 10 × 10 matrices. The superscript k is the number of levels "in between" and ranges from zero to eight. The first four moments of these distributions are given in Table I.

It is perhaps worth noting that an "edge" effect was arbitrarily normalized out of the numerical spacing distribution computations. To understand what this means, consider a 4 × 4 matrix. In the resulting spectrum there are four levels, three nearest-neighbor spacings, two next-nearest-neighbor spacings, and one next-next-nearest-neighbor spacing for each matrix diagonalized. It is obvious that if the mean nearest-neighbor spacing is D, then the mean next-next-nearest-neighbor spacing is $3D$. However, because of the finite size of a 4 × 4 matrix, it is *not* true that the mean next-nearest-neighbor spacing is $2D$ since the two next-nearest-neighbor

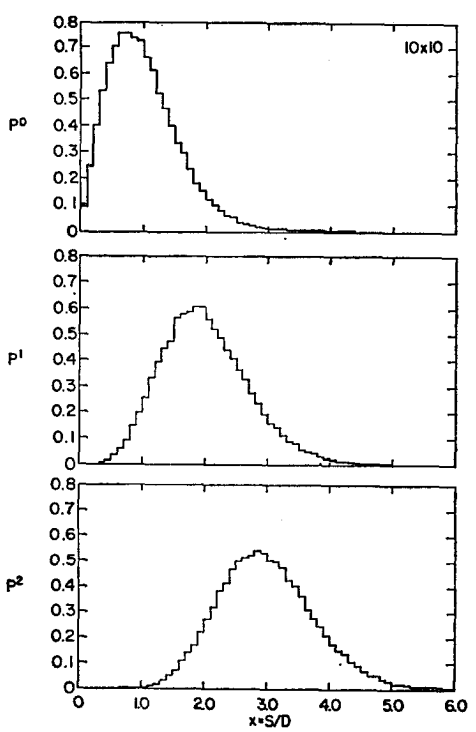

FIG. 1. Histogram plots of the first three (k = 0–2) spacing distributions K^k as a function of $x = S/D$ for 10 × 10 matrices.

[13] K. Fuchel, Rita J. Greibach, and C. E. Porter, "Random Matrix Diagonalization—A Computer Program," Brookhaven National Laboratory Report BNL 760 (T-282), September 1962 (unpublished).

[14] T. J. Krieger and C. E. Porter, J. Math. Phys. (to be published).

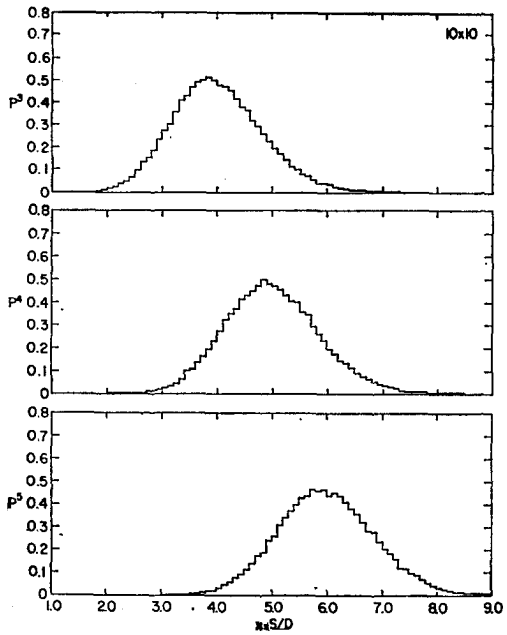

FIG. 2. The same as Fig. 1, except that $k = 3$–5. Note that the abscissa scale does not begin at zero.

spacings do not weight equally the three nearest-neighbor spacings. In the computations, a mean spacing was computed for each P^k and was found to deviate slightly from $(k + 1)D$. This deviation was compensated for by scaling D for each P^k so that the plots in Figs. 1–3 are arranged such that the mean value of x is exactly equal to $k + 1$. This can be seen from Table I; of course, the higher moments are based on the same scale.

Figure 4 shows the sum of all the P^k, i.e.,

$$P^{\text{sum}} = \sum_{k=0}^{N-2} P^k. \tag{3}$$

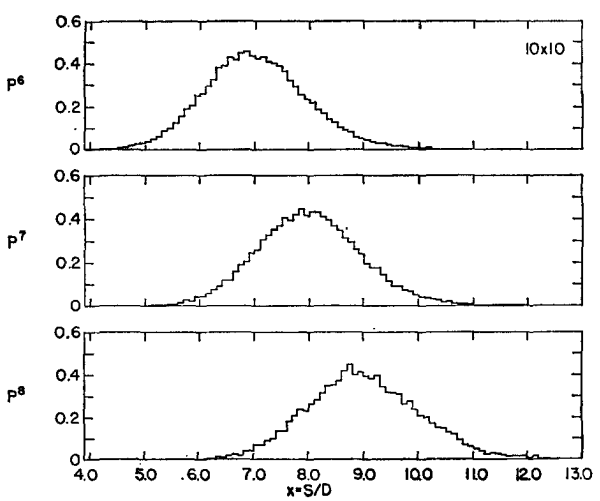

FIG. 3. The same as Fig. 1, except that $k = 6$–8. As in Fig. 2, the abscissa scale does not begin at zero.

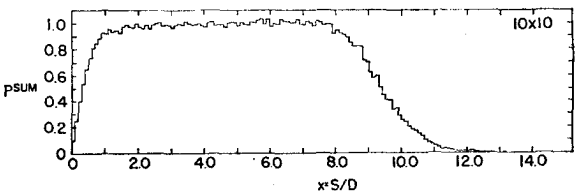

FIG. 4. Histogram plot of P^{sum} as a function of $x = S/D$.

For a Poisson distribution which is appropriate for levels of many different symmetries,

$$P^k(x) = (x^k/k!) \exp(-x), \tag{4}$$

so that in this case

$$P^{\text{sum}}_{\text{Poisson}} \xrightarrow[N\to\infty]{} 1. \tag{5}$$

The repulsion effect for small x is clearly evident in the figure.

Fig. 5 shows the semicircle law of Wigner[15] for the single eigenvalue distribution. The semicircle law holds asymptotically for large N and has the form

$$P(E/2\sigma N^{\frac{1}{2}}) = (2/\pi)[1 - E^2/4\sigma^2 N]^{\frac{1}{2}}. \tag{6}$$

The asymptotic law is clearly already very good for ten dimensions. Analytical expressions for the single eigenvalue distribution for all dimensions in the Gaussian ensemble have been obtained by Mehta and Gaudin.[3] We show the Monte-Carlo results here mainly to indicate the correctness of the computer program.

Because of the plus–minus symmetry of the input matrix element distributions, the first and tenth (in numerical order) eigenvalues are statistically equivalent except for sign. Thus, to increase the counting rate, the sign of the first eigenvalue was changed and the result grouped with the tenth eigenvalue. (We could have likewise plotted only half of the semicircle in Fig. 5 since it is symmetrical about zero.) Similarly the other eigenvalues can be paired off. The resulting histograms are shown in

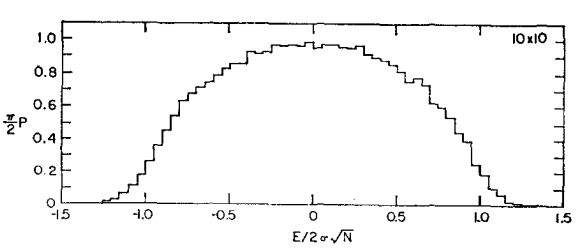

FIG. 5. Histogram plot of the semicircle law. Note that the ordinate is $(\frac{1}{2}\pi)P$ and the abscissa is $E/2\sigma N^{\frac{1}{2}}$.

[15] E. P. Wigner, "Statistical Properties of Real Symmetric Matrices with Many Dimensions" (invited address) Proc. Canadian Math. Congr. (1957), p. 174.

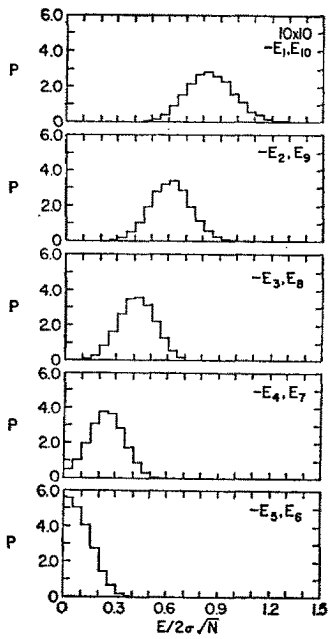

FIG. 6. Histogram plots of the individual eigenvalue distributions as explained in the text. Note that the abscissa is $E/2\sigma N^{\frac{1}{2}}$.

Fig. 6. (Moments of the distributions are given in Table I.) Thus we can see how the semicircle breaks down into its component parts. Another way of saying this is that the separate eigenvalue distributions merge smoothly to form the semicircle, i.e., there is no energy gap in the spectrum. The absence of an energy gap is related to the symmetry of the matrix element distribution.

It is planned to report in the future computations which show that breaking the symmetry of the matrix element distribution leads to an energy gap in the spectrum, with a characteristic departure from the semicircle law.

ACKNOWLEDGMENTS

The computer program for this work was written by Kurt Fuchel and Rita J. Greibach, for which they deserve special thanks. Miss Frances Pope was helpful in plotting some of the preliminary output and also the final results.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 8 AUGUST 1963

Real Unitary Representation of the Many-Channel S Matrix for Complex l and E

CHAN, HONG-MO
Institute for Advanced Study, Princeton, New Jersey
(Received 21 March 1963)

A representation is established for the multichannel S matrix in terms of a matrix function $A(l, E)$ and a scalar function $B(l, E)$, both holomorphic in the domain formed by the product of the whole finite l plane with the finite E plane, cut only along the left-hand dynamical cut. The representation satisfies the known analytic properties of $S(l, E)$, and also all the generalized unitarity conditions of Peierls, LeCouteur, and Newton. The reality condition on S for complex l and E is guaranteed by a simple condition on A and B.

1. INTRODUCTION

THE purpose of this paper is to write down a simple representation of the Schrödinger multichannel S matrix which will automatically satisfy all the unitarity conditions, the reality condition, and the analytic properties of S as a function of the two complex variables l and E. The discussion is mainly kinematical in the sense that it involves only the general characteristics in the formulation of the Schrödinger scattering problem, and does not depend on the details of the interaction. Most of the results are thus expected to hold even in the case of relativistic scattering of particles. For definiteness, however, we may restrict ourselves to the class of potentials studied, for example, by Mandelstam.[1] It was shown[1,2] that the single-channel S matrix is a meromorphic function of the variables l and k in the domain formed by the product of the whole finite l plane, with the k plane cut along the imaginary axis from $k = \frac{1}{2}i\mu$ to $i\infty$, and again

[1] S. Mandelstam, Ann. Phys. (NY), 19, 254, (1962). The potentials studied by Mandelstam are very restrictive. For the more general class studied by Bottino *et al.*, all the discussion in this paper is unchanged except for the domain of analyticity on the l plane, namely the right half-l plane (Re $l > -\frac{1}{2}$) instead of the whole finite l plane.

[2] A. Bottino, A. M. Longoni, and T. Regge, Nuovo Cimento 23, 954 (1962).

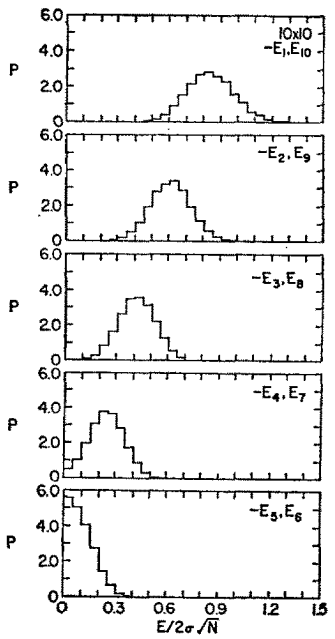

FIG. 6. Histogram plots of the individual eigenvalue distributions as explained in the text. Note that the abscissa is $E/2\sigma N^{\frac{1}{2}}$.

Fig. 6. (Moments of the distributions are given in Table I.) Thus we can see how the semicircle breaks down into its component parts. Another way of saying this is that the separate eigenvalue distributions merge smoothly to form the semicircle, i.e., there is no energy gap in the spectrum. The absence of an energy gap is related to the symmetry of the matrix element distribution.

It is planned to report in the future computations which show that breaking the symmetry of the matrix element distribution leads to an energy gap in the spectrum, with a characteristic departure from the semicircle law.

ACKNOWLEDGMENTS

The computer program for this work was written by Kurt Fuchel and Rita J. Greibach, for which they deserve special thanks. Miss Frances Pope was helpful in plotting some of the preliminary output and also the final results.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 8 AUGUST 1963

Real Unitary Representation of the Many-Channel S Matrix for Complex l and E

CHAN, HONG-MO
Institute for Advanced Study, Princeton, New Jersey
(Received 21 March 1963)

A representation is established for the multichannel S matrix in terms of a matrix function $A(l, E)$ and a scalar function $B(l, E)$, both holomorphic in the domain formed by the product of the whole finite l plane with the finite E plane, cut only along the left-hand dynamical cut. The representation satisfies the known analytic properties of $S(l, E)$, and also all the generalized unitarity conditions of Peierls, LeCouteur, and Newton. The reality condition on S for complex l and E is guaranteed by a simple condition on A and B.

1. INTRODUCTION

THE purpose of this paper is to write down a simple representation of the Schrödinger multichannel S matrix which will automatically satisfy all the unitarity conditions, the reality condition, and the analytic properties of S as a function of the two complex variables l and E. The discussion is mainly kinematical in the sense that it involves only the general characteristics in the formulation of the Schrödinger scattering problem, and does not depend on the details of the interaction. Most of the results are thus expected to hold even in the case of relativistic scattering of particles. For definiteness, however, we may restrict ourselves to the class of potentials studied, for example, by Mandelstam.[1] It was shown[1,2] that the single-channel S matrix is a meromorphic function of the variables l and k in the domain formed by the product of the whole finite l plane, with the k plane cut along the imaginary axis from $k = \frac{1}{2}i\mu$ to $i\infty$, and again

[1] S. Mandelstam, Ann. Phys. (NY), 19, 254, (1962). The potentials studied by Mandelstam are very restrictive. For the more general class studied by Bottino *et al.*, all the discussion in this paper is unchanged except for the domain of analyticity on the l plane, namely the right half-l plane (Re $l > -\frac{1}{2}$) instead of the whole finite l plane.

[2] A. Bottino, A. M. Longoni, and T. Regge, Nuovo Cimento 23, 954 (1962).

from $k = 0$ to $-i\infty$. (Here μ is the inverse range of the potential of interaction.) It is more convenient in what follows to use the energy $E = k^2$ as variable instead of k. $S(l, E)$ is then meromorphic in the product of the finite l plane, with the E plane cut along the real axis from $E = -\frac{1}{4}\mu^2$ to $-\infty$, and again from $E = 0$ to $+\infty$. The branch point at threshold ($E = 0$), studied by Bottino *et al.*,[2] is a "winding" point, and is purely kinematical and independent of the interaction. The cut from $E = -\frac{1}{4}\mu^2$ to $-\infty$, however, depends on the potential and involves the detailed dynamics of the system. In what follows we shall not have occasion to discuss the nature of this "dynamical" cut and our E plane will always be cut from $E = -\frac{1}{4}\mu^2$ to $-\infty$.

By multichannel scattering, we mean the case when the scatterer may have several excited states $a, b, \cdots$, with binding energies $E_a, E_b, \cdots$. The analytic properties of S (i.e. the element of the matrix S) as a function of l have been studied by Charap and Squires[3] and others. For spinless particles, the analytic properties in l are the same as those quoted above for single-channel scattering. The analytic properties of S in the linear momentum for fixed complex l have been studied by Jaffe and Kim[4] in the two-channel case. The analytic properties are similar to the single-channel case. The S matrix considered as a function of k_a, and k_b (where k_a, k_b are the momenta in channels a and b respectively) is meromorphic in the cut $k_a(k_b)$ plane, cut from $k_a(k_b) = \frac{1}{2}i\mu$ to $+i\infty$, and again from $k_a(k_b) = 0$ to $-i\infty$, when k_a and k_b are considered as independent variables. The generalization to more than two channels seems obvious.

The variables $k_a, k_b, \cdots$ are however not independent, but satisfy[5]

$$E = k_a^2 + E_a = k_b^2 + E_b = \cdots . \tag{1.1}$$

Considered as a function of E, S will thus have a series of winding points at $E = E_a, E_b \cdots$, in addition to a left-hand "dynamical" cut along the negative real axis. Otherwise, the function is meromorphic in the whole finite E plane.

The unitarity conditions and related symmetry properties of the multichannel S matrix for integral l have been discussed by several authors.[6-9] Take as illustration the two-channel case, with $E_b > E_a$. From the unitarity condition in the physical region above the threshold of both channels, one can deduce the matrix equation[6]

$$S(k_a, k_b)S(-k_a, -k_b) = 1. \tag{1.2}$$

In the region $E_a < E < E_b$, however, the channel b is not open, and we have a "little" unitarity condition involving channel a only, which may be written as[6]

$$S_{aa}(k_a, k_b)S_{aa}(-k_a, k_b) = 1. \tag{1.3}$$

Note that (1.3) is independent of (1.2). LeCouteur[7] and Newton[8] have found further relations between the matrix elements of S:

$$S_{bb}(k_a, k_b)S_{bb}(k_a, -k_b) = 1, \tag{1.4}$$

$$S_{aa}(k_a, k_b)S_{ab}(-k_a, k_b) = -S_{ab}(k_a, k_b). \tag{1.5}$$

All these symmetries are kinematical and independent of the interaction. Whereas Newton's symmetry (1.5) does not seem to have any direct physical meaning, LeCouteur's relation (1.4) would be the "little" unitarity condition if $E_b < E_a$, and may thus be obtained by analytic continuation in the binding energy E_b from the region $E_b < E_a$ to the region $E_b > E_a$.[10]

These relations (1.2)–(1.5) can be generalized to more than two channels,[9] the number of independent relations increasing as $\sim 2^N$, where N is the number of channels. These conditions relate the values of the different S-matrix elements on different sheets of their Riemann surfaces, and thus represent rather stringent conditions on S as an analytic function. [For example, LeCouteur[7] has shown that for integral l, (1.2)–(1.4) imply that all the matrix elements of S can be generated by a single analytic function defined on the same Riemann surface as that of S.]

For complex l, the structure of the Riemann surface becomes rather complicated and has infinitely many sheets instead of the usual 2^N sheets. The

[3] J. M. Charap and E. J. Squires, Ann. Phys. (NY), **20**, 145 (1962). [I have been informed by J. Charap that the proof quoted here, though valid for spinless particles, is incomplete for particles of arbitrary spins. The proof has been completed in a later paper by the same authors.] I am grateful to J. Charap for this communication.

[4] A. M. Jaffe and Y. S. Kim, Phys. Rev. **127**, 2261 (1962).

[5] One may also easily consider the more general case when the scattered particle had different masses in different channels so that $E = \rho_a k_a^2 + E_a = \rho_b k_b^2 + E_b = \cdots$, where ρ_a, ρ_b are mass ratios for the different channels. We shall ignore this possibility, however, for simplicity.

[6] R. E. Peierls, Proc. Roy. Soc. (London) **A253**, 16, (1959).

[7] E. J. Le Couteur, Proc. Roy. Soc. (London) **A256**, 115 (1960).

[8] R. G. Newton, J. Math. Phys. **2**, 188, (1961).

[9] K. T. R. Davies and M. Baranger, Ann. Phys. (NY), **19**, 383 (1962).

[10] This analytic continuation in the binding energy E_b is immediately available in these simple models of multichannel scattering. Since one knows that S is analytic in k_a and k_b separately when considered as independent variables, Eq. (1.1) then shows that S is analytic in E and E_b separately with E_a fixed.

conditions (1.2)–(1.5) no longer hold but have to be replaced by much more complicated ones. The question was asked whether a simple representation of S can be found which will exhibit all the known analytic properties of S as a function of the two complex variables l and E, and at the same time, satisfy automatically all the symmetry conditions (1.2)–(1.5). Such a representation has been found in terms of a matrix function $A(l, E)$ and a scalar function $B(l, E)$, both holomorphic in the domain formed by the product of the finite l plane with the finite E plane, cut only along the left-hand dynamical cut. The threshold branch points are explicitly exhibited in the representation.

By requiring that $A^*(l, E) = A(l^*, E^*)$ and that a similar condition holds for B, this representation of S will then also satisfy the reality (time-reversal) condition, which, in terms of S, is also rather complicated for complex l because of the winding points at the thresholds.

The representation is closely related to the Wigner[11] R matrix in the theory of nuclear reactions.

The establishment of the representation and the proof that it has the required properties shall be presented in Sec. 3. The procedure is illustrated first for the single-channel case in Sec. 2.

2. REPRESENTATION FOR A SINGLE CHANNEL

Following Bottino *et al.*, we write

$$S(l, k) = [f(l, k)/f(l, ke^{-i\pi})]e^{i\pi l}, \tag{2.1}$$

where $f(l, k)$ is the Jost function for complex l and k. For nonintegral values of l, $f(l, k)$ has a winding point at $k = 0$, such that

$$f(l, ke^{-2i\pi}) = f(l, k) + 2i \sin \pi l f(l, ke^{-i\pi}). \tag{2.2}$$

From time-reversal invariance of the problem, one shows that

$$f^*(l, k) = f(l^*, k^*e^{-i\pi}). \tag{2.3}$$

From (2.1) and (2.3), it is easily shown that

$$S^*(l, k) = S^{-1}(l^*, k^*). \tag{2.4}$$

From (2.1), (2.2), and (2.3), one can also show that

$$S^*(l, k) = S(l^*, k^*e^{i\pi})e^{-2i\pi l^*} + (1 - e^{-2i\pi l^*}), \tag{2.5}$$

and from (2.4),[12]

$$S(l, k)[S(l, ke^{i\pi})e^{-2i\pi l} + (1 - e^{-2i\pi l})] = 1. \tag{2.6}$$

For integral values of l, the conditions (2.5) and (2.6) reduce to the familiar relations

$$S^*(l, k) = S(l, -k^*), \tag{2.7}$$

$$S(l, k)S(l, -k) = 1, \tag{2.8}$$

which are, respectively, the reality condition and the unitarity condition written in analytic form. The conditions (2.5) and (2.6) are thus the correct extensions of the reality and unitarity conditions extended to complex l and k.

Introducing the function $R(l, k)$ as follows,

$$R(l, k) = -ik^{2l+1}[S(l, k) + e^{2i\pi l}]/[S(l, k) - 1], \tag{2.9}$$

we have

$$S(l, k) = [R(l, k) - ik^{2l+1}e^{2i\pi l}]/[R(l, k) + ik^{2l+1}]. \tag{2.10}$$

$R(l, k)$ is the generalization of the Wigner[11] R function to complex values of l and k. It can readily be shown from (2.6)[2] that

$$R(l, ke^{iu\pi}) = R(l, k), \quad u = 0, \pm 1, \pm 2, \cdots, \tag{2.11}$$

i.e., R is an even function of k regular at $k = 0$. With $E = k^2$ as variable, R is thus meromorphic in the domain formed by the product of the finite l plane with the finite E plane, cut only along the left-hand dynamical cut. [Henceforth, this domain shall be designated $D(l, E)$.] Moreover, the converse is also true, i.e. if R is an even function of k, then S defined through (2.10) satisfies the unitarity condition (2.6), as is readily seen by substitution.

Taking the complex conjugate of (2.9), and using (2.5), and (2.11), one finds that the reality condition on R reads

$$R^*(l, k) = e^{-2i\pi l^*}R(l^*, k^*). \tag{2.12}$$

Since R is meromorphic in $D(l, E)$, one can write

$$R(l, E) = e^{i\pi l}[A(l, E)/B(l, E)], \tag{2.13}$$

where (since $e^{i\pi l}$ is entire) A and B are holomorphic functions in $D(l, E)$. For R satisfying (2.12), it shall be shown in Appendix A that one can always find A and B satisfying (2.13), and also[13]

$$A^*(l, E) = A(l^*, E^*); \quad B^*(l, E) = B(l^*, E^*). \tag{2.14}$$

Substituting (2.13) in (2.10), one then has the representation

$$S(l, k) = \frac{A(l, E) - ik^{2l+1}e^{i\pi l}B(l, E)}{A(l, E) + ik^{2l+1}e^{i\pi l}B(l, E)}, \tag{2.15}$$

[11] E. Wigner and L. Eisenbud, Phys. Rev. **72**, 29, (1947).

[12] This formula connects the value of S on one sheet of its Riemann surface to the next and should thus be used for analytic continuation from the physical sheet to other sheets for complex l instead of (2.8), which holds only for integral values of l.

[13] I am grateful to Professor Jost for pointing out that this is not immediately obvious, and also for a subsequent helpful discussion.

where A and B are holomorphic in $D(l, E)$ with $E = k^2$. Moreover, if A and B satisfy (2.14), then the representation satisfies all the known analytic properties of S as a function of l and k, the unitarity condition (2.6) and the reality condition (2.5).

In Sec. 3, we shall generalize (2.15) to multichannel scattering.

3. THE MULTICHANNEL CASE

Following the procedure of Jost and Newton[14] for the case of integral l, we write the Schrödinger equation in matrix form:

$$(d^2/dx^2)\Psi(x) + K^2\Psi(x) - [l(l+1)/x^2]\Psi(x) = V\Psi(x), \quad (3.1)$$

where K is a diagonal matrix, with diagonal elements k_α, and k_α is the momentum in the αth channel. V is a symmetric real matrix potential, where each matrix element satisfies certain conditions, as, e.g., those required for Mandelstam's extension of the Regge formula.[1] We remind the reader that the actual form of the potential is unimportant in this discussion.

We seek matrix solutions $F(l, k, x)$ and $\Phi(l, k, x)$, where $F(l, k, x)$ is defined by the asymptotic condition $F(l, k, x) \to F^0(l, k, x)$ and $\Phi(l, k, x)$ by the boundary condition at the origin $\Phi(l, k, x) \to \Phi^0(l, k, x)$. $F^0(l, k, x)$ and $\Phi^0(l, k, x)$ are both diagonal matrices with

$$F^0_\alpha(l, k, x) = (\tfrac{1}{2}\pi)^{\frac{1}{2}} \times \exp[-i\pi\tfrac{1}{2}(l+1)](k_\alpha x)^{\frac{1}{2}}H^{(2)}_{l+\frac{1}{2}}(k_\alpha x), \quad (3.2)$$

$$\Phi^0_\alpha(l, k, x) = 2^{l+\frac{1}{2}}\Gamma(l+\tfrac{3}{2})x^{\frac{1}{2}}k_\alpha^{-l-\frac{1}{2}}J_{l+\frac{1}{2}}(k_\alpha x). \quad (3.3)$$

These solutions, $F(l, k, x)$ and $\Phi(l, k, x)$ may be defined uniquely by integral equations as in Bottino *et al.*[2] We shall write down only the equation for $F(l, k, x)$:

$$F(l, k, x) = F^0(l, k, x) + \int_x^\infty G(l, k, x, \xi)V(\xi)F(l, k, \xi)\,d\xi. \quad (3.4)$$

The Green-function matrix $G(l, k, x, \xi)$ is a diagonal matrix with elements

$$G_\alpha(l, k, x, \xi) = i^{\frac{1}{4}}\pi x^{\frac{1}{2}}\xi^{\frac{1}{2}}[H^{(1)}_{l+\frac{1}{2}}(k_\alpha\xi)H^{(2)}_{l+\frac{1}{2}}(k_\alpha x) - H^{(2)}_{l+\frac{1}{2}}(k_\alpha\xi)H^{(1)}_{l+\frac{1}{2}}(k_\alpha x)]. \quad (3.5)$$

One may then define Jost matrix as the Wronskian

$$F(l, k) = W[\Phi(l, k, x), F(l, k, x)], \quad (3.6)$$

and the S matrix, following Bottino *et al.*[2,4] as

$$S(l, k) = K^{\frac{1}{2}}F^{-1}(l, ke^{-i\pi})F(l, k)K^{-\frac{1}{2}}e^{i\pi l}. \quad (3.7)$$

[14] See, e.g., R. G. Newton, J. Math. Phys. **1**, 319 (1960).

In the above formulas, we have used k to denote loosely the set of variables k_α, $\alpha = a, b, c \cdots$. Thus, e.g., by $k \to ke^{-i\pi}$ in (3.7), we mean $k_\alpha \to k_\alpha e^{-i\pi}$ for all α. It must be remembered, however, that these variables k_α are not independent but are related by (1.1), so that quantities like the Jost matrix $F(l, k)$ and the S matrix $S(l, k)$ are functions of only one momentum (energy) variable which may be taken as E. Because of (1.1), and the fact that their definition depend explicitly on k_α, these quantities, as functions of E, will be multivalued with branch points at $E = E_\alpha$, $\alpha = a, b, c, \cdots$, in addition to any further branch points they may have as functions of the k_α's. It is thus convenient to keep the notation with k explicit (interdependence understood) in order to specify which branch of the functions we are considering. Thus, for example, $k_\alpha \to k_\alpha e^{-i\pi}$ with all other k_β's fixed, means that we take a circuit on the E plane once clockwise around the point $E = E_\alpha$ without enclosing any other branch point; while $k \to k\,e^{-i\pi}$ in (3.7), means we take a circuit clockwise once around all the branch points $E = E_\alpha$, $\alpha = a, b, \cdots$.

We shall now study the branching properties of the quantities defined above in order to establish the required representation of the S matrix. From the property of the Henkel function and the definition (3.2), one sees that[2]

$$F^0_\alpha(l, k_\alpha e^{-2i\pi}, x) = F^0_\alpha(l, k_\alpha, x) + 2i\sin\pi l F^0_\alpha(l, k_\alpha e^{-i\pi}, x). \quad (3.8)$$

Since $F^0_\beta(l, k_\beta, x)$ depends only on k_β and not on any other k_α ($\alpha \neq \beta$), the circuit $k_\alpha \to k_\alpha e^{-2i\pi}$ will leave all other $F^0_\beta(l, k_\beta, x)$ unchanged. This may be summarized in the notation

$$F^0(l, k, x; k_\alpha \to k_\alpha e^{-2i\pi}) = F^0(l, k, x) + 2i\sin\pi l F^0(l, k, x; k_\alpha \to k_\alpha e^{-i\pi})P_\alpha, \quad (3.9)$$

where P_α is the projection matrix into the αth channel, namely a diagonal matrix with unity as the αth element, and all other elements zero.

From the fact that $G(l, k, x, \xi)$ is regular at $k_\alpha = 0$, $\alpha = a, b, c \cdots$, and the circuit relation (3.9) of $F^0(l, k, x)$, it can be seen by substitution that

$$F(l, k, x) + 2i\sin\pi l F(l, k, x; k_\alpha \to k_\alpha e^{-i\pi})P_\alpha$$

satisfies the equation for $F(l, k, x; k_\alpha \to k_\alpha e^{-2i\pi})$. The solution $F(l, k, x)$ thus satisfies the same circuit law (3.9) as $F^0(l, k, x)$. Notice, however, that $F(l, k, x)$ is not in general diagonal, and the presence of the projection operator P_α means that only the

αth column is altered by the double circuit around $E = E_\alpha$, all other columns of $F(l, k, x)$ remaining unchanged.

It can readily be seen that the other solution $\Phi(l, k, x)$ is regular at $k_\alpha = 0$, so that the definition (3.6) and the circuit law (3.9) for $F(l, k, x)$ imply that the Jost matrix $F(l, k)$ satisfies the same circuit law (3.9).

A similar investigation as above shows that in fact the βth column of $F(l, k)$ is even under a change of sign of k_α, $\alpha \neq \beta$, i.e., more precisely,

$$F(l, k; k_\alpha \to k_\alpha e^{-i\pi})P_\beta = F(l, k)P_\beta, \quad \beta \neq \alpha. \tag{3.10}$$

From the circuit laws (3.9) and (3.10) of the Jost matrix, one can deduce the branching properties of the S matrix at the thresholds $E = E_\alpha$ from the definition (3.7). One may deduce relations between the matrix elements of $S(l, k)$ and the matrix elements of S with the sign of a number of k's changed. These relations would be generalizations of the relation (2.6) to several channels, and also analytic continuations to complex l of the generalized unitarity conditions of Peierls–LeCouteur–Newton.[6-9] These relations are, however, very complicated and are, in fact, what we set out to avoid, and since they are, in any case, implicitly contained in the representation that follows, they are no longer discussed.

Introduce next the Wigner R matrix in a similar way to (2.9) as

$$R = -iK^{l+\frac{1}{2}}[S - 1]^{-1}[S + e^{2i\pi l}]K^{l+\frac{1}{2}}. \tag{3.11}$$

We wish to show, as in Sec. 2, that R is in fact meromorphic in $D(l, E)$. By definition (3.11), R is meromorphic whenever S is meromorphic. The individual terms on the right-hand side of (3.11) have, however, branch points at $E = E_\alpha$, $\alpha = a, b, c, \cdots$. To show that R is meromorphic in $D(l, E)$, it is thus sufficient and necessary to show that these branch points disappear in the combination (3.11). We must, therefore, show that R is unchanged by a small circuit around $E = E_\alpha$ (i.e. $k_\alpha \to k_\alpha e^{-i\pi}$) for any α, i.e., the infinitely many sheets of the Riemann surface all collapse into the E plane.

By definition (3.7), R may be written

$$R(l, k) = -iK^{l+1}[F(l, k)e^{i\pi l} - F(l, ke^{-i\pi})]^{-1} \times [F(l, k)e^{i\pi l} + F(l, ke^{-i\pi})e^{2i\pi l}]K^l. \tag{3.12}$$

Consider now the circuit $k_\alpha \to k_\alpha e^{-i\pi}$, with all other k_β unchanged. Then $K^l \to K^l C_\alpha^l$, where C_α is a diagonal matrix with $e^{-i\pi}$ as the αth element and all other elements unity. Thus,

$$C_\alpha = (1 - P_\alpha) + e^{-i\pi}P_\alpha. \tag{3.13}$$

Also, using the circuit law (3.9),

$$\begin{aligned} F(l, k; k_\alpha \to k_\alpha e^{-i\pi})e^{i\pi l} &- F(l, ke^{-i\pi}; k_\alpha \to k_\alpha e^{-i\pi}) \\ &= F(l, k; k_\alpha \to k_\alpha e^{-i\pi})e^{i\pi l} \\ &\quad - F(l, ke^{-i\pi}; k_\alpha \to k_\alpha e^{i\pi}) \\ &\quad - 2i \sin \pi l F(l, ke^{-i\pi})P_\alpha. \end{aligned} \tag{3.14}$$

By circuit law (3.10), this may be rewritten as

$$\begin{aligned} [F(l, k)e^{i\pi l} &- F(l, ke^{-i\pi})][(1 - P_\alpha) + e^{-i\pi(l+1)}P_\alpha] \\ &= [F(l, k)e^{i\pi l} - F(l, ke^{-i\pi})]C^{l+1}. \end{aligned} \tag{3.15}$$

Similarly, it can be seen that

$$\begin{aligned} F(l, k; k_\alpha \to k_\alpha e^{-i\pi})e^{i\pi l} &\\ + F(l, ke^{-i\pi}; k_\alpha \to k_\alpha e^{-i\pi})e^{2i\pi l} &\\ = [F(l, k)e^{i\pi l} + F(l, ke^{-i\pi})e^{2i\pi l}]C^{-l}. \end{aligned} \tag{3.16}$$

The factors C^{l+1} in (3.15) and C^l in (3.16) just cancel those from the K factors outside in the expression (3.12). R is thus invariant under the circuit $k_\alpha \to k_\alpha e^{-i\pi}$ for any α. The meromorphy of R in $D(l, E)$ is then established.

We may then write each of the matrix elements in the form (2.13), where, without loss of generality, we may take B to be the same for all elements. Substituting into the definition (3.11) and solving for S, one has

$$\begin{aligned} S = K^{-(l+\frac{1}{2})}&(A - iK^{2l+1}e^{-i\pi l}B) \\ &\times (A + iK^{2l+1}e^{-i\pi l}B)^{-1}K^{l+\frac{1}{2}}. \end{aligned} \tag{3.17}$$

This is the desired representation. $A = A(l, E)$ is a matrix, and $B = B(l, E)$ a scalar. Both A and B are holomorphic in the domain $D(l, E)$. If A and B satisfy (2.14), then the reality condition of S is satisfied by the representation (3.17). That (3.17) satisfies all the generalized unitarity conditions of Peierls–Le Couteur–Newton[6-9] is a matter only of algebra involving elementary properties of determinants, and will be proved in Appendix B.

We have thus proved (a) that the representation (3.17) is always valid with A and B holomorphic in $D(l, E)$, and (b) that the representation satisfies the reality condition, the generalized unitarity con ditions, and all the analytic properties required. Such a representation might be useful for further investigation of the many-channel S matrix. For example, we see that Regge trajectories are given by a equation of the form

$$\det (A + iK^{2l+1}e^{-i\pi l}B) = 0, \tag{3.18}$$

which shows the manner in which a Regge trajectory behaves at threshold, and as it crosses the threshold cuts into other sheets.

ACKNOWLEDGMENT

The author would like to thank Professor J. R. Oppenheimer for hospitality at the Institute for Advanced Study, and to thank many members of the Institute, especially Professors F. Dyson, R. Jost, and N. N. Khuri for helpful discussions, and for reading the manuscript.

APPENDIX A

We shall show that if R satisfies (2.12), then A and B holomorphic in $D(l, E)$ may be found satisfying both (2.13) and (2.14).

One may assume without loss of generality that A and B in (2.13) do not have common zeros. Then (2.12) implies

$$A^*(l, E) = C(l^*, E^*)A(l^*, E^*),$$
$$B^*(l, E) = C(l^*, E^*)B(l^*, E^*), \qquad (A.1)$$

where $C(l^*, E^*)$ is any function of l^*, E^*. Since $A(l, E)$ holomorphic in l, E implies $A^*(l, E)$ holomorphic in l^*, E^*, (A.1) requires that $C(l^*, E^*)$ is meromorphic in l^*, E^* in $D(l, E)$. (D is symmetric under the operation $l, E \rightarrow l^*, E^*$.) Since, by hypothesis, A and B have no common zeros, $C(l, E)$ by (A.1) has neither zeros nor poles in $D(l, E)$, i.e., $C(l, E)$ is entire. Taking complex conjugate of (A.1),

$$A(l, E) = C^*(l^*, E^*)C(l^*, E^*)A(l, E), \qquad (A.2)$$

we have

$$C^*(l^*, E^*)C(l^*, E^*) = 1, \qquad (A.3)$$

except possibly where $A(l, E) = 0$. A being holomorphic, can only have isolated zeros, and C must then satisfy

$$C^*(l, E)C(l, E) = 1 \qquad (A.4)$$

throughout D, by analyticity.

Since $C(l, E)$ is entire, $C^{\frac{1}{2}}(l, E)$ is also entire. One may then define two new functions

$$A'(l, E) = [C(l, E)]^{\frac{1}{2}}A(l, E),$$
$$B'(l, E) = [C(l, E)]^{\frac{1}{2}}B(l, E), \qquad (A.5)$$

which will then satisfy both (2.12) and (2.13), and remain holomorphic in $D(l, E)$.

APPENDIX B

The generalized unitarity conditions of Peierls–Le Couteur–Newton[6-9] may be best stated in the general case in the form of Davies and Baranger.[9] Let the many-channel S matrix be partitioned into submatrices

$$S = \begin{pmatrix} S_{PP} & S_{PQ} \\ S_{QP} & S_{QQ} \end{pmatrix}.$$

Then

$$S_{PP}(k_P, k_Q)S_{PP}(-k_P, k_Q) = 1, \qquad (B.1)$$
$$S_{PP}(k_P, k_Q)S_{PQ}(-k_P, k_Q) = (-1)^{l+\frac{1}{2}}S_{PQ}(k_P, k_Q). \qquad (B.2)$$

Consider two matrices $C = (C_{ij})$, $D = (d_{ij})$. Let $C^* = (C_{ij}^*)$ be the matrix formed by replacing the first m rows of C by the corresponding rows of D, and similarly for $D^* = (d_{ij}^*)$. Thus,

$$C_{ij}^* = d_{ij}, \qquad d_{ij}^* = C_{ij},$$

for

$$i = 1, \cdots, m, \qquad j = 1, \cdots, N.$$

Then $D^{-1} = (\bar{D}_{ji}/\Delta)$, where $\Delta = \det(d_{ij})$, $\bar{D}_{ji}$ is the minor of d_{ij} in D. This gives, for $A = CD^{-1}$, $a_{ij} = N_{ij}/\Delta$, where $N_{ij} = \sum_k C_{ik}\bar{D}_{jk}$.

We wish to show that

$$\sum_{i=1}^{m} a_{ji}a_{ij}^* = 1, \quad \text{for} \quad j = 1, \cdots, m.$$

This follows, since with $C_{ik} = d_{ik}^*$, $C_{il}^* = d_{il}$,

$$\sum_{i=1}^{m} N_{ji}N_{ij}^* = \sum_{l=1}^{N} d_{jl}^*\bar{D}_{jl}^* \sum_{i=1}^{m} d_{il}\bar{D}_{il} + \sum_{k\neq l} d_{jk}^*\bar{D}_{jl}^* \sum_{i=1}^{m} d_{il}\bar{D}_{ik}$$
$$= \Delta^*\left(\Delta - \sum_{p=m+1}^{N} d_{pl}\bar{D}_{pk}\right) + \sum_{k\neq l} d_{jk}^*\bar{D}_{jl}^*\left(-\sum_{p=m+1}^{N} d_{pl}\bar{D}_{pk}\right)$$
$$= \Delta^*\Delta - \sum_{p=m+1}^{N} \sum_{k} d_{jk}^*\bar{D}_{pk} \sum_{l} d_{pl}\bar{D}_{jl}^*,$$

the last term being zero since $d_{pl} = d_{pl}^*$ for $p = m + 1$ and $j \neq p$.

Similarly, it can be seen that

$$\sum_{i=1}^{m} a_{ji}a_{ik}^* = \delta_{jk} \quad \text{for} \quad j, k \leq m, \qquad (B.3)$$

and

$$\sum_{i=1}^{m} a_{ji}a_{ip}^* = -a_{jp} \quad \text{for} \quad j \leq m, p > m. \qquad (B.4)$$

It only remains to be noted that the many-channel S matrix for integral l both in the form (3.7) and in the representation (3.17) is of the form $C(k)C^{-1}(-k)$ apart from factors of the diagonal matrix K. Inverting the signs of a number of k variables corresponds to the * operation above. Using the relations (B.3) and (B.4), it is then easy to demonstrate that the conditions (B.1) and (B.2) are satisfied for both (3.7) and (3.17) for integral values of l.

Relativistic Potential Scattering*

REESE T. PROSSER
Lincoln Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received 16 April 1963)

The scattering properties of the relativistic two-body problem, governed by the Dirac equation, are investigated. It is shown rigorously that the associated Hamiltonians are self-adjoint, that the associated wave operators exist, and that the scattering operator exists and is unitary, all under suitable conditions on the potential. These conditions on the potential are analogues of those required for the nonrelativistic two-body problem governed by the Schrödinger equation.

INTRODUCTION

IN recent years there has appeared a substantial body of literature devoted to a mathematical formulation of scattering theory. Most of this work has focused on the nonrelativistic two-body problem governed by the Schrödinger equation. Results obtained so far include a proof of the self-adjointness of the associated Hamiltonians,[1] a proof of the existence of the associated wave operators,[2,3] and a proof of the existence and unitary property of the scattering operator,[4,5] all under suitable conditions on the potential. The concepts involved here appear to be fundamental in any approach to scattering theory, and we may regard these results as a major contribution toward a rigorous description of the nonrelativistic two-body problem.[6]

It is tempting to suppose that the methods which have proved so successful for the two-body problem might be used to obtain a rigorous description of some of the many-body problems of field theory. These problems, however, are all cast in relativistic form, and require an appropriate modification of non-relativistic techniques. As a first step toward these problems, we present here a study of the relativistic two-body problem governed by the Dirac equation. We show that the associated Hamiltonians are self-adjoint, that the associated wave operators exist, and that the scattering operator exists and is unitary, all under suitable conditions on the potential. The conditions required are given explicitly in the statements of the theorems below; they turn out to be the analogues of those required in the nonrelativistic case.

* Operated with support from the U. S. Army, Navy, and Air Force.

[1] T. Kato, Trans. Am. Math. Soc. **70**, 195 (1951).
[2] J. M. Cook, J. Math. Phys. **36**, 82 (1957).
[3] S. T. Kuroda, J. Math. Soc. Japan **11**, 247 (1959).
[4] T. Ikebe, Arch. Rational Mech. Anal. **5**, 1 (1960).
[5] R. T. Prosser, "A Convergent Perturbation Expansion for Wave Operators," J. Math. Phys. (to be published).
[6] S. T. Kuroda, Nuovo Cimento **12**, 431 (1959).

1. GENERAL FRAMEWORK

We shall deal throughout the paper with the following general framework: Let $\mathfrak{H}$ denote a separable Hilbert space, and $\mathfrak{D}$ a fixed dense domain in $\mathfrak{H}$. Let H_0 denote a symmetric operator defined on $\mathfrak{D}$ which is essentially self-adjoint there.[1] Let V denote a symmetric operator defined on $\mathfrak{D}$ and majorized by H_0 in a suitable sense, to be made precise below. Put

$$H = H_0 + V. \tag{1.1}$$

In this situation we have at our disposal a number of quite general theorems describing the properties of H in terms of those of H_0 and V. We shall state here without proof those which are essential to our purpose. The first of them, established in 1951 by Kato,[1] runs as follows:

Theorem 1.1. *Suppose for each* $f \in \mathfrak{D}$, V *satisfies*

$$||Vf|| \leq \alpha\,||H_0 f|| + \beta\,||f||, \tag{1.2}$$

where α *and* β *are positive scalars independent of* f, *and* $\alpha < 1$. *Then* H *is essentially self-adjoint on* $\mathfrak{D}$.

Proof: See Kato.[1]

If H is essentially self-adjoint on $\mathfrak{D}$, then we may form the one-parameter unitary groups $\exp(iH_0t)$ and $\exp(iHt)$ and define the *transition operators*

$$U(s, t) = \exp(-iH_0 s) \times \exp[iH(s - t)] \exp(iH_0 t). \tag{1.3}$$

In terms of the transition operators, we may define the *incoming* and *outgoing wave operators*

$$W_{\mp} = \lim_{t \to \mp\infty} U(0, t), \tag{1.4}$$

whenever these limits exist in a suitable sense. Our next result, which is due to Cook,[2] gives a criterion for the existence of these limits in the strong topology of $\mathfrak{H}$. The criterion is expressed in terms of a one-parameter family of potentials derived from V,

$$V(t) = \exp(-iH_0 t)V \exp(iH_0 t). \tag{1.5}$$

Theorem 1.2. *Suppose that for each f in a dense subset of $\mathfrak{D}$ we have*

$$\int_{-\infty}^{+\infty} ||V(t)f||\, dt < \infty.$$

Then the wave operators $W_\pm$ of (1.4) *are defined in the strong topology of $\mathfrak{H}$, and $HW_\pm = W_\pm H_0$.*

Proof: See Cook,[2] or Kuroda.[3]

Finally, we may define the *scattering operator*

$$S = \lim_{s\to+\infty} \lim_{t\to-\infty} U(s, t), \tag{1.6}$$

whenever these limits exist. Our final result gives a sufficient condition for the existence of these limits in the strong topology of $\mathfrak{H}$.

Theorem 1.3. *Suppose V can be expressed as a product AB of two operators A and B defined on $\mathfrak{D}$ and satisfying*

(a) *A is bounded,*

(b) *$B \exp(iH_0t)Af$ is defined for each $f \in \mathfrak{D}$, and $||B \exp(iH_0t)Af|| \leq K(t)\, ||f||$, where $K(t)$ is independent of f,*

(c) $$\int_{-\infty}^{+\infty} K(t)\, dt < 1. \tag{1.7}$$

Then the scattering operator S given by (1.6) *is defined in the strong topology of $\mathfrak{H}$, and is essentially unitary there. Moreover, $H_0S = SH_0$.*

Proof: See Prosser.[5]

For a discussion of the properties and physical interpretation of the wave and scattering operators, we refer the reader to references 6 and 7.

2. RELATIVISTIC FREE-PARTICLE WAVE FUNCTIONS

According to relativistic quantum theory, the wave function associated with a single relativistic free electron satisfies the Dirac equation

$$(i\gamma\partial + m)\psi(x) = 0. \tag{2.1}$$

[We are using here the notation of reference 7: ψ is a four-component spinor-valued function on the Minkowski space–time manifold M_4, γ denotes the Dirac matrix vector, and ∂ the space–time gradient. The scalar m represents the particle mass in suitably normalized units. We adapt for the Lorentz metric on M_4 the signature $(+, -, -, -)$.]

Every (sufficiently regular) solution of (2.1) is the Fourier transform of a spinor-valued measure μ;

$$\psi(x) = (2\pi)^{-2} \int_{M_4} e^{i(px)}\, d\mu(p). \tag{2.2}$$

It follows from (2.1) that $\mu(p)$ must satisfy

$$(\gamma p - m)\mu(p) = 0. \tag{2.3}$$

This system of four equations admits nontrivial solutions only if

$$\det(\gamma p - m) = p^2 - m^2 = 0. \tag{2.4}$$

Thus, $\mu(p)$ must have support on the hyperboloid $p^2 - m^2 = 0$. We shall consider here only measures of the form

$$\mu(p) = \delta(p^2 - m^2)\chi(p), \tag{2.5}$$

where $\chi(p)$ is a spinor-valued function which is locally integrable on M_4.

Not all measures of the form (2.5), however, are solutions of (2.3). To see this, we introduce the matrix-valued functions $\Lambda_\pm(p)$,

$$\Lambda_\pm(p) = (\pm\gamma p + m)/2m, \tag{2.6}$$

which satisfy the conditions

$$\begin{aligned} \Lambda_\pm(p)^2 &= \Lambda_\pm(p), \\ \Lambda_\pm(p)\Lambda_\mp(p) &= (-p^2 + m^2)/4m^2, \\ \Lambda_+(p) + \Lambda_-(p) &= 1, \\ \Lambda_-(+p) &= \Lambda_+(-p). \end{aligned} \tag{2.7}$$

We then define the projection operators $\Lambda_\pm$ on measures of the form (2.5) by

$$(\Lambda_\pm\mu)(p) = \delta(p^2 - m^2)\Lambda_\pm(p)\chi(p). \tag{2.8}$$

Then we have

$$\mu = \Lambda_+\mu + \Lambda_-\mu, \tag{2.9}$$

and it follows from (2.7) that μ satisfies (2.4) if and only if $\Lambda_-\mu = 0$, and in this case we have

$$\begin{aligned} \mu(p) &= (\Lambda_+\mu)(p) \\ &= \delta(p^2 - m^2)[(\gamma p + m)/(2m)]\chi(p). \end{aligned} \tag{2.10}$$

Substituting (2.10) into (2.2) and integrating over p_0, we obtain

$$\psi(x) = (2\pi)^{-\frac{3}{2}} \int [e^{ipx}\chi^+(p) + e^{-ipx}\chi^-(p)]\, d\mathbf{p}/\epsilon(\mathbf{p}), \tag{2.11}$$

where we have put

$$p_0 = \epsilon(\mathbf{p}) = +(\mathbf{p}^2 + m^2)^{\frac{1}{2}}, \tag{2.12}$$

and

$$\chi^\pm(p) = (8\pi)^{-\frac{1}{2}}\Lambda_\pm(p)\chi(\pm p). \tag{2.13}$$

These considerations suggest the following arrangement: Let $\mathfrak{H}$ denote the space of all spinor-

[7] S. S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Row, Peterson and Company, Evanston, Illinois, 1961).

valued functions $\psi(p)$ defined on E_3 such that the norm

$$||\psi||^2 = \int_{E_3} |\psi(\mathbf{p})|^2 \, d\mathbf{p} \tag{2.14}$$

is finite. Here, $|\psi(\mathbf{p})|^2$ denotes the sum of the squares of the absolute values of the four components of $\psi(\mathbf{p})$. $\mathfrak{H}$ (modulo functions of norm zero) forms a Hilbert space under this norm. For each function $\psi(\mathbf{p})$ in $\mathfrak{H}$, define

$$\psi^{\pm}(\mathbf{p}) = \pm[\gamma^0 m/\epsilon(\mathbf{p})]\Lambda_{\pm}(\pm\mathbf{p})\psi(\mathbf{p}), \tag{2.15}$$

where $\Lambda_{\pm}(\mathbf{p})$ is the function $\Lambda_{\pm}(p)$ with $p_0 = \epsilon(\mathbf{p})$. We observe that $\psi^+(\mathbf{p})$ and $\psi^-(\mathbf{p})$ are orthogonal in $\mathfrak{H}$, and

$$\psi(\mathbf{p}) = \psi^+(\mathbf{p}) + \psi^-(\mathbf{p}). \tag{2.16}$$

Hence,

$$||\psi||^2 = ||\psi^+||^2 + ||\psi^-||^2. \tag{2.17}$$

Let $\mathfrak{D}$ be the subspace of $\mathfrak{H}$ consisting of those functions $\psi(\mathbf{p})$ such that $\epsilon(\mathbf{p})\psi(\mathbf{p})$ also lies in $\mathfrak{H}$, and for such functions define

$$(H_0\psi)(\mathbf{p}) = p_0\psi(\mathbf{p}) = \epsilon(\mathbf{p})\psi^+(\mathbf{p}) - \epsilon(\mathbf{p})\psi^-(\mathbf{p}). \tag{2.18}$$

Then, since H_0 is a multiplication operator acting on a maximal domain, it is well-defined and self-adjoint on $\mathfrak{D}$. The action of e^{iH_0t} on $\psi(\mathbf{p})$ is given by

$$(e^{iH_0t}\psi)(\mathbf{p}) = e^{i\epsilon(\mathbf{p})t}\psi^+(\mathbf{p}) + e^{-i\epsilon(\mathbf{p})t}\psi^-(\mathbf{p}). \tag{2.19}$$

Next, define $\psi(\mathbf{x})$ via

$$\psi(\mathbf{x}) = (2\pi)^{-\frac{3}{2}} \int_{E_3} e^{i\mathbf{p}\mathbf{x}}\psi(\mathbf{p}) \, d\mathbf{p}, \tag{2.20}$$

and observe that from the Plancherel relations we have

$$||\psi||^2 = \int_{E_3} |\psi(\mathbf{x})|^2 \, d\mathbf{x}. \tag{2.21}$$

Thus, we may regard $\psi(\mathbf{p})$ and $\psi(\mathbf{x})$ as different representations of the same element ψ of $\mathfrak{H}$. The actions of H_0 and e^{iH_0t} on $\psi(\mathbf{x})$ are determined via (2.20); in particular, we have

$$(e^{iH_0t}\psi)(\mathbf{x}) = (2\pi)^{-\frac{3}{2}} \times \int_{E_3} e^{i\mathbf{p}\mathbf{x}}[e^{i\epsilon(\mathbf{p})t}\psi^+(\mathbf{p}) + e^{-i\epsilon(\mathbf{p})t}\psi^-(\mathbf{p})] \, d\mathbf{p}. \tag{2.22}$$

This agrees with (2.11) if we put

$$\chi^{\pm}(p) = \pm\epsilon(\mathbf{p})\psi^{\pm}(\pm\mathbf{p}). \tag{2.23}$$

It follows that $\psi(\mathbf{x}, t) = (e^{iH_0t}\psi)(\mathbf{x})$ is a (weak) solution of the Dirac equation (2.1), and that every solution which lies in $\mathfrak{H}$ at time $t = 0$ must be of this form. In particular, the development in time of every such solution is unitary. (More generally, it can be shown that the norm (2.21) is invariant under the action induced on $\mathfrak{H}$ by the inhomogeneous Lorentz group, and that this action constitutes an irreducible representation of the group.)[8]

Next we present an explicit description of the action of e^{iH_0t} on $\psi(\mathbf{x})$.

Theorem 2.1. *Suppose $\psi(\mathbf{x})$ is continuously differentiable on E_3. Then for each $t \neq 0$, the function $(e^{iH_0t}\psi)(\mathbf{x})$ is given by*

$$e^{iH_0t}\psi(\mathbf{x}) = iS_t * \gamma^0\psi(\mathbf{x}) = (2\pi)^{-\frac{3}{2}} \int_{E_3} iS(\mathbf{y})\gamma^0\psi(\mathbf{x} - \mathbf{y}) \, d\mathbf{y} \tag{2.24}$$

where $S_t(\mathbf{x})$ is a matrix-valued distribution on E_3 given by

$$S_t(\mathbf{x}) = S(\mathbf{x}, t) = S^+(\mathbf{x}, t) + S^-(\mathbf{x}, t). \tag{2.25}$$

Here

$$S^{\pm}(\mathbf{x}, t) = (-i\gamma\partial + m)\Delta^{\pm}(\mathbf{x}, t), \tag{2.26}$$

with

$$\Delta^+(\mathbf{x}, t) = (2\pi)^{\frac{1}{2}}m^2/2 \times \begin{cases} -\dfrac{H_1^{(2)}[m(t^2 - r^2)^{\frac{1}{2}}]}{m(t^2 - r^2)^{\frac{1}{2}}} & \text{if} \quad t < -r, \\ \dfrac{2}{\pi}\dfrac{K_1[m(r^2 - t^2)^{\frac{1}{2}}]}{m(r^2 - t^2)^{\frac{1}{2}}} & \text{if} \quad -r < t < r, \\ +\dfrac{H_1^{(1)}[m(t^2 - r^2)^{\frac{1}{2}}]}{m(t^2 - r^2)^{\frac{1}{2}}} & \text{if} \quad r < t, \end{cases} \tag{2.27}$$

and $\Delta^-(\mathbf{x}, t) = \Delta^+(\mathbf{x}, t)^*$. The symbol $*$ in (2.24) denotes convolution on E_3, to be interpreted as described below.

Proof: From (2.15) and (2.22) we have

$$(e^{iH_0t}\psi)(\mathbf{x}) = (e^{iH_0t}\psi^+)(\mathbf{x}) + (e^{iH_0t}\psi^-)(\mathbf{x}), \tag{2.28}$$

where

$$(e^{iH_0t}\psi^{\pm})(\mathbf{x}) = (2\pi)^{-\frac{3}{2}} \times \int_{E_3} e^{i\mathbf{p}\mathbf{x}}e^{\pm i\epsilon(\mathbf{p})t}\frac{\gamma^0 m}{\epsilon(\mathbf{p})}\Lambda_{\pm}(\pm\mathbf{p})\psi(\mathbf{p}) \, d\mathbf{p}. \tag{2.29}$$

Now consider the effect of replacing t by $t_1 = t + i\delta$, where $\delta > 0$. Then $(e^{iH_0t_1}\psi^+)(\mathbf{x})$ is well-defined via (2.29), and we have

$$e^{iH_0t}\psi^+ = iS^+ * \gamma^0\psi = (2\pi)^{-\frac{3}{2}} \int_{E_3} iS^+(\mathbf{y})\gamma^0\psi(\mathbf{x} - \mathbf{y}) \, d\mathbf{y}, \tag{2.30}$$

[8] V. Bargmann and E. P. Wigner, Proc. Natl. Acad. Sci. U. S. 34, 211 (1948).

where

$$S^+(\mathbf{x}) = S^+(\mathbf{x}, t_1) = (2\pi)^{-\frac{3}{2}} \int_{E_3} e^{i\mathbf{p}\mathbf{x}} e^{i\epsilon(\mathbf{p})t_1} 2m\Lambda_+(\mathbf{p}) \frac{d\mathbf{p}}{\epsilon(\mathbf{p})} = (-i\gamma\partial + m)(2\pi)^{-\frac{3}{2}} \int_{E_3} e^{i\mathbf{p}\mathbf{x}} e^{i\epsilon(\mathbf{p})t_1} \frac{d\mathbf{p}}{\epsilon(\mathbf{p})} \quad (2.31)$$
$$= (-i\gamma\partial + m)\Delta^+(\mathbf{x}, t_1).$$

Note that as long as $\delta > 0$, $\Delta^+(\mathbf{x}, t_1)$, and hence $S^+(\mathbf{x}, t_1)$, is an integrable real analytic function of $\mathbf{x}$ on E_3, so that the convolution in (2.30) is well defined. Similarly, $(e^{iH_0t_1^*}\psi^-)(\mathbf{x})$ is well defined and given by

$$e^{iH_0t_1^*}\psi = iS^- * \gamma^0\psi, \quad (2.32)$$

where

$$S^-(\mathbf{x}) = S^-(\mathbf{x}, t_1^*) = (-i\gamma\partial + m)\Delta^-(\mathbf{x}, t_1^*). \quad (2.33)$$

Now a standard bounded convergence argument, together with the Plancherel relations, shows that $iS^+ * \gamma^0\psi + iS^- * \gamma^0\psi$ converges in the mean-square sense to $e^{iH_0t}\psi$ as $\delta \to 0$.

Now the function $\Delta(x, t_1) = \Delta^+(\mathbf{x}, t_1) + \Delta^-(\mathbf{x}, t_1^*)$ remains analytic on E_3 as $\delta \to 0$, except on the sphere $t^2 - r^2 = 0$, where we find a singularity of the form

$$\Delta(\mathbf{x}, t) = (2\pi)^{\frac{1}{2}}\eta(t) \times [\delta(t^2 - r^2) - \tfrac{1}{2}m^2\theta(t^2 - r^2) + \cdots]. \quad (2.34)$$

Here the remainder is continuous at $t^2 = r^2$. $\theta(\xi)$ is the standard Heaviside step function given by

$$\theta(\xi) = \begin{cases} 1 & \text{if } \xi \geq 0, \\ 0 & \text{if } \xi < 0, \end{cases} \quad (2.35)$$

and $\eta(\xi)$ is given by

$$\eta(\xi) = 2\theta(\xi) - 1. \quad (2.36)$$

Moreover, $\Delta(\mathbf{x}, t)$ vanishes outside this sphere. A similar statement holds, of course, for $S(\mathbf{x}, t)$.

Now if ψ is continuously differentiable on E_3, then $\psi(\mathbf{x}-\mathbf{y})$ may be expanded near the sphere $t^2 - r^2$ in the form $[\mathbf{x} - \mathbf{y} = (r, \theta, \varphi)]$,

$$\psi(\mathbf{x} - \mathbf{y}) = A(\theta, \varphi) + (t^2 - r^2)B(r, \theta, \varphi), \quad (2.37)$$

where $B(r, \theta, \varphi)$ is continuous at $t^2 = r^2$. It follows by standard methods that the limit as $\delta \to 0$ of the convolution (2.24) may be evaluated for each point $\mathbf{x}$ of E_3.

Of special interest is a variant of theorem 2.1 which holds whenever $\psi(\mathbf{x})$ has compact support and t is so large that the sphere $t^2 - r^2 = 0$ lies outside this support. Then the singularities of S_t do not enter the convolution in (2.24) for small values of $\mathbf{x}$.

To formulate this result precisely, we introduce the projection operator E_R on $\mathfrak{H}$ by

$$(E_R\psi)(\mathbf{x}) = \begin{cases} \psi(x) & \text{if } |x| \leq R, \\ 0 & \text{if } |x| > R. \end{cases} \quad (2.38)$$

Corollary 2.2. *Suppose* $\psi \in \mathfrak{H}$ *and* $t > 2R$. *Then*

$$(E_Re^{iH_0t}E_R\psi)(\mathbf{x}) = i(E_{2R}S_t) * \gamma^0(E_R\psi)(\mathbf{x}). \quad (2.39)$$

Proof: If ψ is continuously differentiable, we have

$$(E_Re^{iH_0t}E_R\psi)(\mathbf{x}) = iE_R(S_t * \gamma^0E_R\psi) = \begin{cases} (2\pi)^{-\frac{3}{2}} \int_{E_3} iS(\mathbf{y})\gamma^0(E_R\psi)(\mathbf{x} - \mathbf{y})\, d\mathbf{y} & \text{if } |x| \leq R, \\ 0 & \text{if } |x| > R. \end{cases} \quad (2.40)$$

Now if $|\mathbf{x}| \leq R$ and $|\mathbf{x} - \mathbf{y}| \leq R$, then $|\mathbf{y}| \leq 2R$, and the integral in (2.40) vanishes for $|\mathbf{y}| > 2R$. Hence we may replace $S_t(\mathbf{y})$ by $(E_{2R}S_t)(\mathbf{y})$ in the convolution, and obtain (2.39) for differentiable functions. Now if $t > 2R$, then $(E_{2R}S_t)(\mathbf{y})$ has no singularities in E_3, and the convolution in (2.39) is well defined for all ψ in $\mathfrak{H}$. Since differentiable functions are dense in $\mathfrak{H}$, we conclude that (2.39) holds for all ψ in $\mathfrak{H}$, as required.

3. INTRODUCTION OF POTENTIALS

In the presence of a static electromagnetic field, the Dirac equation (2.1) is modified by the addition of a term describing the action of the field potential

$$(i\gamma\partial + m)\psi(x) = eV(x)\gamma^0\psi(x), \quad (3.1)$$

where $V(x)$ is a prescribed function representing the field potential, and e is a positive scalar representing the charge. We shall assume throughout that $V(x) = V(\mathbf{x})$ is independent of time, and is locally integrable on E_3.

If $\psi(x)$ is any solution of (3.1), then

$$(-i\partial_0\psi)(x) = \gamma^0[i\gamma\partial - m + eV(\mathbf{x})\gamma^0]\psi(x) = (H_0 + V)\psi(\mathbf{x}) \quad (3.2)$$

where $H_0\psi$ is defined in (2.18), and $V\psi$ is given by

$$(V\psi)(\mathbf{x}) = eV(\mathbf{x})\psi(\mathbf{x}). \quad (3.3)$$

It follows that if $\psi(\mathbf{x})$ lies in H at time $t = x_0 = 0$, then it lies in $\mathfrak{H}$ for each time t, and its development in time is given by

$$\psi(x) = \psi(\mathbf{x}, t) = e^{iHt}\psi(\mathbf{x}), \quad (3.4)$$

provided only that $H = H_0 + V$ is well-defined and essentially self-adjoint on $\mathfrak{H}$.

Now V is a multiplication operator on $\psi(\mathbf{x})$. If $V(\mathbf{x})$ is bounded, then V is a bounded operator, and satisfies the condition (1.2) of theorem 1.1 with $\alpha = 0$. It follows from theorem 1.1 that $\mathfrak{H}$ is essentially self-adjoint on $\mathfrak{D}$.

It is of some interest to extend this result to unbounded potentials. Our next result gives conditions on $V(\mathbf{x})$ which ensure that theorem 1.1 applies even through $V(\mathbf{x})$ is unbounded. We utilize the arguments developed by Kato[1] and extended by Brownell.[9]

Theorem 3.1. *Suppose* $V(\mathbf{x})$ *lies in* $\mathfrak{L}_p$, *where* $3 < p \leq \infty$. *Then*

$$||V\psi||_2 \leq \alpha\, ||H_0\psi||_2 + \beta\, ||\psi||_2, \tag{3.5}$$

with $\alpha < 1$.

Proof: According to a straightforward extension of the Hölder inequality,[10] we have

$$||V\psi||_2 \leq ||V||_p\, ||\psi||_q, \tag{3.6}$$

where $1/p + 1/q = \frac{1}{2}$, and $p > 3$, $q > 6$. Now a similar extension of the Hausdorff–Young inequality[11] yields

$$||\psi||_q \leq \text{const}\, ||\chi||_r, \tag{3.7}$$

where $1/q + 1/r = 1$, with $q > 6$, $r < \frac{6}{5}$, and χ is related to ψ via (2.11). Here we define the $\mathfrak{L}_r$ norm of χ by

$$||\chi||_r^r = \int_{E_*} |\chi(\mathbf{p})|^r\, d\mathbf{p}. \tag{3.8}$$

Finally, the extended Hölder inequality gives

$$||\chi||_r \leq ||(\epsilon + \lambda)\chi||_2\, ||(\epsilon + \lambda)^{-1}||_s, \tag{3.9}$$

where $\frac{1}{2} + 1/s = 1/r$, with $r < \frac{6}{5}$, $s > 3$. Here $\epsilon + \lambda$ denotes the function $\epsilon(\mathbf{p}) + \lambda$, with λ a positive constant at our disposal.

Now we note

$$\begin{aligned} ||(\epsilon + \lambda)\chi||_2 &\leq ||\epsilon\chi||_2 + \lambda\, ||\chi||_2 \\ &= ||H_0\chi||_2 + \lambda\, ||\chi||_2, \end{aligned} \tag{3.10}$$

which is finite whenever $\chi \in \mathfrak{D}$. Moreover,

$$||(\epsilon + \lambda)^{-1}||_s \leq \text{const}\, \lambda^{(3-s)/s}, \tag{3.11}$$

since $s > 3$. Thus if $\chi \in \mathfrak{D}$, we have

$$||\chi||_r \leq \text{const}\, (\lambda^{-1}\, ||H_0\chi||_2 + ||\chi||_2)\lambda^{3/s} \tag{3.12}$$

and it follows from the Plancherel relations that

$$\begin{aligned} ||V\psi||_2 &\leq ||V||_p\, ||\psi||_q \\ &\leq \text{const}\, ||V||_p\, (\lambda^{-1}\, ||H_0\psi||_2 + ||\psi||_2)\lambda^{3/s} \end{aligned} \tag{3.13}$$

whenever $\psi \in \mathfrak{D}$. Since λ is arbitrary, we conclude that the operation of multiplication by $V(\mathbf{x})$ is majorized by H_0 in the sense of (3.5) with α arbitrarily small. Theorem 1.1 now tells us that $H = H_0 + V$ is essentially self-adjoint on $\mathfrak{D}$.

If $V(\mathbf{x})$ is the sum of potentials in different $\mathfrak{L}_p$ spaces ($p > 3$), then, since each summand satisfies (3.5) with α arbitrarily small, so does V. In particular, if $V(\mathbf{x})$ is a sum of potentials with disjoint supports (apart from sets of measure zero) and lying in distinct $\mathfrak{L}_p$ spaces (p always > 3), then V satisfies (3.5) and theorem 1.1 holds.

It is unfortunate that theorem 3.1 does not cover the Coulomb and Yukawa potentials [$V(\mathbf{x}) \sim (4\pi r)^{-1}\, e^{-\mu r}$] because the singularity at the origin is not p-summable for $p > 3$. Apparently a deeper analysis is required for these special cases. We shall not attempt to include them here.

4. EXISTENCE OF THE WAVE OPERATORS

We have now constructed the framework described in Sec. 1. In this section we shall show that the hypotheses of Theorems 1.2 and 1.3 are satisfied within this framework.

We begin with a lemma describing the behavior of functions of the $(E_R\, e^{iH_0t}E_R\psi)(\mathbf{x})$ as $t \to \pm\infty$, where E_R is given by (2.38).

Lemma 4.1. *Suppose* $\psi \in \mathfrak{H}$. *Then*

$$||E_R e^{iH_0t}E_R\psi||_2 \leq ||E_R\psi||_2 \qquad \text{for all } t, \tag{4.1}$$

$$||E_R e^{iH_0t}E_R\psi||_\infty \leq \text{const}\, |t|^{-\frac{3}{2}}\, ||E_R\psi||_1 \text{ for } |t| \geq 4R. \tag{4.2}$$

Proof: The first of these follows from the fact that e^{iH_0t} is unitary and E_R is a projection, and it holds for all times t. For the second, we use corollary 2.2 and observe that

$$\begin{aligned} ||E_R e^{iH_0t}E_R\psi||_\infty &= ||(E_{2R}S_t) * \gamma^0(E_R\psi)||_\infty \\ &\leq ||E_{2R}S_t||_\infty\, ||E_R\psi||_1. \end{aligned} \tag{4.3}$$

Now if ψ is in $\mathfrak{H}$, then $E_R\psi$ is integrable, and $||E_R\psi||_1$ is finite. On the other hand, if $|t| > 2R$, then $E_{2R}S_t$ is bounded, since

$$E_{2R}S_t(\mathbf{x}) = \begin{cases} (-i\gamma\partial + m)\Delta(\mathbf{x}, t) & \text{if } |\mathbf{x}| \leq 2R, \\ 0 & \text{if } |\mathbf{x}| > 2R, \end{cases} \tag{4.4}$$

and $\Delta(\mathbf{x}, t)$ is singular only where $|x| = |t|$. Now if $|\mathbf{x}| < 2R < |t|$, then

[9] F. H. Brownell, Pac. J. Math. **12**, 47 (1962).

[10] N. Dunford, and J. Schwarz, *Linear Operators* (Interscience Publishers, Inc., New York, 1958), p. 527.

[11] E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals* (Clarendon Press, Oxford, England, 1937), p. 96.

$$(i\gamma\partial - m)\Delta(\mathbf{x}, t) = \text{const}\,(i\gamma\partial - m)\frac{J_1[m(t^2 - r^2)^{\frac{1}{2}}]}{m(t^2 - r^2)^{\frac{1}{2}}}$$

$$= \text{const}\left\{\frac{J_2[m(t^2 - r^2)^{\frac{1}{2}}]}{[m(t^2 - r^2)^{\frac{1}{2}}]^2}\, i(\gamma^0 t + \boldsymbol{\gamma}\mathbf{x}) - m\,\frac{J_1[m(t^2 - r^2)^{\frac{1}{2}}]}{m(t^2 - r^2)^{\frac{1}{2}}}\right\}. \quad (4.5)$$

This function is certainly bounded for $|r| < 2R$. Moreover, it follows from known properties of the Bessel functions[12] that, for $r < 2R < 4R \leq |t|$,

$$\left|\frac{J_p[m(t^2 - r^2)^{\frac{1}{2}}]}{[m(t^2 - r^2)^{\frac{1}{2}}]^p}\right| \leq \text{const}\,(t^2 - r^2)^{-\frac{1}{4}(2p+1)} \leq \text{const}\,|t|^{-\frac{1}{2}(2p+1)}. \quad (4.6)$$

On the other hand, the term $|\gamma^0 t + \boldsymbol{\gamma}\mathbf{x}|$ is majorized by const $|t|$. Combining, we find that

$$||E_{2R}S_t(x)||_\infty \leq \text{const}\,|t|^{-\frac{3}{2}}, \quad (4.7)$$

and (4.2) follows from (4.3) and (4.7). Finally, we observe that the constant in (4.2) may be chosen independent of R as well as of t.

Now we turn to the question of existence of the wave operators. We shall show that under appropriate conditions on the potential $V(\mathbf{x})$, the conditions of theorem 1.2 are satisfied.

Theorem 4.2. *Suppose $V(\mathbf{x})$ is such that*

(a) $||E_R V||_2 < \infty$ *for all* R,

(b) $||(1 - E_R)V||_\infty \leq \text{const}\ R^{-\eta}$ *for large* R *and some* $\eta > 1$.

Then if $\psi(\mathbf{x})$ is infinitely differentiable and has compact support,

$$\int_{-\infty}^{+\infty} ||V(t)\psi||_2\, dt < \infty, \quad (4.8)$$

and H_0 and V satisfy the conditions of Theorem 1.2.

Proof: Choose R_0 large enough so that both $E_{R_0}\psi = \psi$ and $||(I - E_R)V||_\infty \leq \text{const}\ R^{-\eta}$ for all $R > R_0$. Now suppose $|t| = 4R > 4R_0$, and consider

$$||V(t)\psi||_2 = ||Ve^{iH_0 t}\psi||_2$$
$$\leq ||E_{R_0}Ve^{iH_0 t}\psi||_2 + ||(E_R - E_{R_0})Ve^{iH_0 t}\psi||_2$$
$$+ ||(I - E_R)Ve^{iH_0 t}\psi||_2. \quad (4.9)$$

The first term here is majorized by

$$||E_{R_0}Ve^{iH_0 t}\psi||_2 = ||VE_{R_0}e^{iH_0 t}E_{R_0}\psi||_2$$
$$\leq ||E_{R_0}V||_2\, ||E_{R_0}e^{iH_0 t}E_{R_0}\psi||_\infty$$
$$\leq \text{const}\,|t|^{-\frac{3}{2}}, \quad \text{since} \quad |t| > 4R_0. \quad (4.10)$$

The second term of (4.9) is majorized by

$$||(E_R - E_{R_0})Ve^{iH_0 t}\psi||_2$$
$$\leq ||(E_R - E_{R_0})V||_2\, ||E_R e^{iH_0 t}E_R\psi||_\infty$$
$$\leq \text{const}\,R^{\frac{1}{2}(3-2\eta)} \times \text{const}\,|t|^{-\frac{3}{2}}$$
$$\leq \text{const}\,|t|^{-\eta}. \quad (4.11)$$

Finally, the last term in (4.9) is majorized by

$$||(I - E_R)Ve^{iH_0 t}\psi||_2 \leq ||(I - E_R)V||_\infty\, ||e^{iH_0 t}\psi||_2$$
$$\leq \text{const}\,R^{-\eta}\, ||\psi||_2$$
$$\leq \text{const}\,|t|^{-\eta}. \quad (4.12)$$

Thus we conclude that if $|t| > 4R_0$, then

$$||V(t)\psi||_2 \leq \text{const}\,t^{-\eta}. \quad (4.13)$$

Now if $\psi(\mathbf{x})$ is infinitely differentiable and has compact support on E_3, then it follows from (2.24) that the same is true of $\psi(\mathbf{x}, t)$ for each time t. Thus for $|t| < 4R_0$ the function $\psi(\mathbf{x}, t)$ is infinitely differentiable and has compact support on M_4, and in particular is bounded there. It follows that for $|t| < 4R_0$,

$$||V(t)\psi||_2 = ||E_{5R_0}Ve^{iH_0 t}\psi||_2$$
$$\leq ||E_{5R_0}V||_2\, ||e^{iH_0 t}\psi||_\infty$$
$$\leq \text{const}\,||\psi||_2. \quad (4.14)$$

Combining (4.13) and (4.14), we obtain (4.8), as required.

In a quite similar way, we shall establish the existence and unitarity of the scattering operator by showing that, under appropriate restrictions on the potential, the conditions of Theorem 1.3 are satisfied.

Theorem 4.3. *Suppose $V(\mathbf{x})$ is bounded and has compact support. Choose R large enough so that $E_R V = V$. Then*

$$||E_R e^{iH_0 t}V\psi||_2 \leq K(t)\, ||\psi||_2, \quad (4.15)$$

where

$$\int_{-\infty}^{+\infty} K(t)\, dt = \text{const}\,||V||_\infty. \quad (4.16)$$

In particular, H_0 and V satisfy the conditions of Theorem 1.3 *if $||V||_\infty$ is sufficiently small.*

Proof: If $|t| < 4R$, we have

$$||E_R e^{iH_0 t}V\psi||_2 \leq ||V\psi||_2 \leq ||V||_\infty\, ||\psi||_2. \quad (4.17)$$

If $|t| \geq 4R$, then by Lemma 4.1,

$$||E_R e^{iH_0 t}V\psi||_2 = ||E_R e^{iH_0 t}E_R V\psi||_2$$
$$\leq \text{const}\,|t|^{-\frac{3}{2}}\, ||V\psi||_1$$
$$\leq \text{const}\,|t|^{-\frac{3}{2}}\, ||E_R V||_2\, ||\psi||_2$$
$$\leq \text{const}\,|t|^{-\frac{3}{2}}\tfrac{4}{3}\pi R^3\, ||V||_\infty\, ||\psi||_2. \quad (4.18)$$

[12] G. N. Watson, *Treatise on the Theory of Bessel Functions*, (Cambridge University Press, Cambridge, England, 1922).

Combining (4.17) and (4.18), we obtain (4.15), as required.

5. DISCUSSION

We have shown that, under suitable conditions on the potential, the Hamiltonian operator for the relativistic two-body problem is essentially self-adjoint on the associated state space and that the wave operators and the scattering operator exist in the strong sense. The conditions required for these results are more or less natural analogues of those required in the nonrelativistic form of the problem.

There are one or two differences, however, which should be emphasized. We have already observed that our proof of the self-adjointness of H (Theorem 3.1) excludes the Coulomb and Yukawa potentials because of their singularity at the origin. This is not true in the nonrelativistic case, and the difference is a reflection of the fundamental difference between the relativistic and nonrelativistic free-particle Hamiltonian operators. We note here, however, that our proof of the existence of the wave operators does not exclude the Yukawa potentials, so that if the self-adjointness of the Hamiltonian is known for these potentials, then the existence of the wave operators is assured. The Coulomb potential remains excluded because of its behavior at infinity, just as in the nonrelativistic case.

The conditions required for our proof of the existence and unitarity of the scattering operator are too restrictive to admit any potentials of physical interest. Moreover, they imply that the system admits no bound states.[5] It seems likely that our formulation of this result is not the best possible for the relativistic two-body problem, and that the scattering operator actually exists for a much broader class of potentials. The best formulation of this result in the nonrelativistic case comes from a theorem of Ikebe[4] on the existence of an eigenfunction expansion for the total Hamiltonian of the problem. The scattering operator may be described explicitly in terms of these eigenfunctions, which must in some sense contain all the structure inherent in the Hamiltonian. It seems reasonable to suppose that a similar result can be obtained in this way for the relativistic case.

In spite of its shortcomings, our result commands a certain academic interest. First of all, it does prove that Dyson's construction of the scattering operator can be made rigorous at least for a class of potentials broad enough to approximate any physically realizable potential as closely as desired. Second, our result forms a starting point for the investigation of certain problems of quantum field theory, in which the interaction terms contain form factors which approximate the delta functions of point interactions, and which may be conveniently chosen to satisfy the requirements of Theorem 4.3.

There is no difficulty in principle in extending the results of this paper to potentials with several components. In particular, the tensor and spin-orbit potentials of nuclear scattering problems may be treated by the same methods.

Continuous-Representation Theory. I. Postulates of Continuous-Representation Theory

JOHN R. KLAUDER
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received 26 December 1962)

In a continuous representation of Hilbert space, each vector Ψ is represented by a complex, continuous, bounded function $\psi(\Phi) \equiv (\Phi, \Psi)$ defined on a set $\mathfrak{S}$ of continuously many, nonindependent unit vectors Φ having rather special properties: Each vector in $\mathfrak{S}$ possesses an arbitrarily close neighboring vector, and the identity operator is expressable as an integral over projections onto individual vectors in $\mathfrak{S}$. In particular cases it is convenient to introduce labels for the vectors in $\mathfrak{S}$ whereupon each Ψ is represented by a complex, continuous, bounded, label-space function. Basic properties common to all continuous representations are presented, and some applications of the general formalism are indicated.

INTRODUCTION

THE essential ingredients in the basic structure of quantum mechanics are surprisingly few: (i) unit vectors in a Hilbert space $\mathfrak{H}$ correspond to states of a system; (ii) dynamics or scattering involves an automorphism among unit vectors; and (iii) the natural inner products in $\mathfrak{H}$ are interpreted as probability amplitudes. Physics enters, as well as being read out of the formalism, by means of one or more mappings $\mathfrak{M}$ from appropriate label sets characterizing the physical problem into unit vectors in $\mathfrak{H}$. For example, the "in" and "out" states of scattering theory are manifestations of two different mappings into $\mathfrak{H}$ of a parameter set including four-momentum, spin, baryon number, etc. However, independent of the particular form a mapping takes, it provides the "bridge" between the abstract quantum formalism and the label-space framework in which stochastic statements pertinent to a particular system are made.

Now dynamics, or any other Lie group of automorphisms, entails a continuous permutation among unit vectors in $\mathfrak{H}$. It would frequently be desirable to have Hilbert-space representations expressed as label-space functions, admitting direct parameterization of such continuous transformations simply by means of label-space transformations. That is, the labels by themselves are rich enough to parameterize continuous permutations among unit vectors in $\mathfrak{H}$. This clearly requires that the lables must in part assume values in the continuum. For present purposes, such a requirement rules out representations wherein vectors are functions of discrete variables, e.g., the eigenvalues of a complete set of commuting observables with discrete spectrum. The Dirac prescription to generate representations with continuous labels (e.g., the Schrödinger representation) cannot be regarded as a continuous mapping of the label space into unit vectors in $\mathfrak{H}$. Instead, continuity demands that the image set of unit vectors $\mathfrak{S}$ cannot be an orthonormal set, but, in contrast, $\mathfrak{S}$ *must contain vectors arbitrarily close to one another*. This basic property is common to all the image sets $\mathfrak{S}$ we consider, and is fundamental for the definition of a continuous representation. Accounts of specific label sets, mappings, and continuous representations will be treated in the following paper,[1] and in subsequent papers in connection with various applications. Here it is our purpose to present the requirements on the mapping $\mathfrak{M}$ and the image sets $\mathfrak{S}$ that are necessary for a continuous representation to exist, and to discuss some of the basic properties of such representations common to all systems.

POSTULATES OF CONTINUOUS-REPRESENTATION THEORY

We choose a Hilbert space $\mathfrak{H}$, finite- or infinite-dimensional, with positive-definite metric. Among all the vectors in $\mathfrak{H}$, let us focus our attention on *unit* vectors, to be universally denoted by Φ with an arbitrary array of sub- or superscripts, etc. Thus

$$(\Phi, \Phi)^{\frac{1}{2}} \equiv ||\Phi|| = 1 = ||\Phi_1|| = ||\Phi'|| = \cdots .$$

Let $\mathfrak{T}$ denote the set of all unit vectors, and $\mathfrak{S}$ denote a subset of $\mathfrak{T}$: $\mathfrak{S} \subset \mathfrak{T}$. It is the set $\mathfrak{S}$ that will be the image set of the mapping $\mathfrak{M}$. Then, we assert

Postulate 1. (Local density and continuity). For each $\Phi \in \mathfrak{S}$ and every $\delta > 0$, there exists a vector $\Phi' \in \mathfrak{S}$, different from Φ, for which $||\Phi - \Phi'|| < \delta$. The set $\mathfrak{S}$ is an arcwise connected subset of $\mathfrak{H}$ or a union thereof.

[1] J. R. Klauder, J. Math. Phys. 4, 1058 (1963) (following paper).

Clearly, the conventional orthonormal basis set fails to satisfy Postulate 1.

Further, let $\mathfrak{L}$ denote the "label" space whose points l may be correlated with vectors in $\mathfrak{S}$ by the mapping $\mathfrak{M}$ of $\mathfrak{L}$ onto $\mathfrak{S}$. Regarding this correlation, we require

Postulate 2. (Label continuity). The mapping $\mathfrak{M}$: $l \to \Phi[l]$ is a many-one continuous map of a Hausdorff, i.e., separable topological space $\mathfrak{L}$ onto $\mathfrak{S}$. By continuity in $\mathfrak{S}$, we mean the usual weak continuity in $\mathfrak{H}$. Thus if $l_n \to l$, then $(\Phi[l_n], \Psi) \to (\Phi[l], \Psi)$ for all $\Psi \in \mathfrak{H}$.

It is often the case that the topology for $\mathfrak{L}$ is the usual topology where open sets are identified as open intervals. The basic purpose of this postulate is to provide a parameterization, i.e., a "handle" for the states in $\mathfrak{S}$.

For purposes of using the vectors in $\mathfrak{S}$ to define a representation with as familiar a form as possible, we assume

Postulate 3. (Completeness and resolution). The set $\mathfrak{S}$ spans the space $\mathfrak{H}$, i.e., completion in norm of the set of all linear combinations of elements in $\mathfrak{S}$ yields $\mathfrak{H}$. A resolution of unity in $\mathfrak{H}$ exists as an integral over projection operators onto individual vectors in $\mathfrak{S}$.

When $\mathfrak{S}$ is locally compact, then the last postulate means that some additive real measure μ on elementary sets in $\mathfrak{S}$ exists such that

$$\Psi = \int_{\mathfrak{S}} \Phi \, d\mu\,(\Phi)(\Phi, \Psi),$$

or

$$(\Psi', \Psi) = \int_{\mathfrak{S}} (\Psi', \Phi)\, d\mu\,(\Phi)(\Phi, \Psi), \tag{1}$$

for arbitrary vectors Ψ and Ψ'. If $\mathfrak{S}$ is not locally compact, the concept of integral can be generalized to give meaning to (1), e.g., in the manner of Friedrichs of Shapiro.[2] The only role of μ shall be to generate the expansion (1), a restriction which, in general, is insufficient to fix μ uniquely. It is certainly plausible, therefore, and we shall provide a simple proof in the Appendix, for a compact set $\mathfrak{S}$, that (i) because of invariance under unitary transformations, and (ii) the existence of (1), we have

Theorem 1. (Invariant measure.) There is no loss of generality in assuming μ invariant under any and all unitary transformations U that leave the compact set $\mathfrak{S}$ invariant: $U\mathfrak{S} = \mathfrak{S}$. The stated invariance takes the form $\mu(U\mathfrak{R}) = \mu(\mathfrak{R})$ for all $\mathfrak{R} \subset \mathfrak{S}$. A suitable choice for μ is one for which $d\mu(U\Phi) = d\mu(\Phi)$. Consequently, an important application of Theorem 1 arises if the group $\mathcal{G}$ of such unitary transformations U forms a transitive permutation group on vectors in $\mathfrak{S}$ (i.e., if $\Phi \in \mathfrak{S}$, then $\{U\Phi \mid U \in \mathcal{G}\} = \mathfrak{S}$), for then μ can be chosen without loss of generality as the invariant group measure. Examples of this approach to find μ are given in the following paper.[1]

Form of the Continuous Representation

Equation (1) provides the basis for a continuous representation of Hilbert space. In such a representation, the vector Ψ is represented by the complex, bounded, continuous function

$$\psi(\Phi) \equiv (\Phi, \Psi); \qquad \psi^*(\Phi) \equiv (\Psi, \Phi), \tag{2}$$

[or by $\psi(l) \equiv (\Phi[l], \Psi)$ if specific labels are introduced]. The inner product of two vectors is then a restatement of (1):

$$(\Psi', \Psi) = \int_{\mathfrak{S}} \psi'^*(\Phi)\, d\mu\,(\Phi)\psi(\Phi). \tag{3}$$

Not all functions on $\mathfrak{S}$ represent vectors, but only those which satisfy

$$\psi(\Phi') = \int_{\mathfrak{S}} \mathcal{K}(\Phi'; \Phi)\, d\mu\,(\Phi)\psi(\Phi), \tag{4}$$

where the reproducing kernel

$$\mathcal{K}(\Phi'; \Phi) \equiv (\Phi', \Phi), \tag{5}$$

as follows from Eqs. (3) and (2). Furthermore, $\mathcal{K}$ fulfills the indempotent relation

$$\mathcal{K}(\Phi'; \Phi'') = \int_{\mathfrak{S}} \mathcal{K}(\Phi'; \Phi)\, d\mu\,(\Phi)\mathcal{K}(\Phi; \Phi''). \tag{6}$$

An operator $\mathcal{B}$ defined on $\mathfrak{S}$ is represented as a function of two points in $\mathfrak{S}$ by

$$\mathcal{B}(\Phi'; \Phi) \equiv (\Phi', \mathcal{B}\Phi) \tag{7}$$

[or by $\mathcal{B}(l'; l)$ if specific labels are introduced], which is separately continuous in each argument and bounded if $\mathcal{B}$ is a bounded operator. The representation of $\mathcal{B}\Psi$ is clearly

$$(\mathcal{B}\psi)(\Phi') = \int_{\mathfrak{S}} \mathcal{B}(\Phi'; \Phi)\, d\mu\,(\Phi)\psi(\Phi). \tag{8}$$

Not every function of two points in $\mathfrak{S}$ represents an operator, but only those which satisfy

[2] K. O. Friedricks and H. N. Shapiro, Proc. Natl. Acad. Sci. U. S. 43, 336 (1957); *Integration of Functionals*, New York University, Institute of Mathematical Sciences (1957). Here we treat such cases formally.

$$\mathcal{B}(\Phi'; \Phi'') = \int_{\mathfrak{S}} \mathcal{B}(\Phi'; \Phi)\, d\mu\,(\Phi)\mathcal{K}(\Phi; \Phi'')$$

$$= \int_{\mathfrak{S}} \mathcal{K}(\Phi'; \Phi)\, d\mu\,(\Phi)\mathcal{B}(\Phi; \Phi'')$$

$$= \iint_{\mathfrak{S}} \mathcal{K}(\Phi'; \Phi_1)\, d\mu\,(\Phi_1)$$

$$\times \mathcal{B}(\Phi_1; \Phi_2)\, d\mu\,(\Phi_2)\mathcal{K}(\Phi_2; \Phi''), \quad (9)$$

as follows from (7) and (8) and their adjoints.

It is clear that the set $\mathfrak{S}$ plays an important role in establishing a continuous representation and determining its properties: Postulate 1 leads, with the aid of continuity in the inner product, to continuous functions ψ defined on unit vectors; Postulate 2 leads, in turn, to continuous functions ψ defined on points of label space; and Postulate 3 ensures that the functions ψ provide a representation of $\mathfrak{H}$. As a matter of nomenclature, we call an abstract set of unit vectors $\mathfrak{S}$ that satisfies Postulates 1 and 3 an *overcomplete family of states* (OFS). As is common in quantum mechanics, we shall often refer to the continuous representation generated (in the manner described above) by an OFS simply by reference to the particular OFS itself.

Some Applications of Overcomplete Families of States

An OFS may be used as any other representation would be used in quantum or quantum statistical mechanics. As such, the OFS appears generally to be a more proper way to introduce vectors with continuous labels than the conventionally used but nonexistent eigenstates for operators with continuous spectra. In particular, there exists one special choice of an OFS that is closely related to the Fock representation of an infinite-dimensional Hilbert space by entire analytic functions, which has recently seen a renewed interest.[3] For applications to field theories, one natural set of "labels" for the OFS are the well-defined test functions of Distribution Theory.[4]

In the following paper[1] we shall establish a generalized form of "classical" dynamics expressed in terms of the continuously variable c-number labels that characterize vectors of the OFS. This analysis shows, for example, that, for certain classes of Hamiltonians, we can exactly describe quantum mechanics by the classical dynamical formalism merely by reinterpreting the classical dynamical variables as c-number labels belonging to Hilbert space vectors. In addition, it is possible that overcomplete families of states may prove useful in placing the sum-over-histories on a more sound mathematical footing; already in formal studies they have permitted the usual Fresnel integrals to be replaced by absolutely convergent Gaussian integrals.[5] On the other hand, an OFS may be used for a Hamiltonian-*less* approach to dynamics directly through a postulated evolutionary[6] or in–out scattering automorphism.

Finally, it should be remarked that a natural choice of an OFS for fermion degrees of freedom[7] corresponds to the set of states generated by all Bogoliubov transformations. Thus, this and other OFS may be directly relevant in making interesting approximations.

The author thanks J. McKenna for several discussions.

APPENDIX

We wish to prove for compact sets $\mathfrak{S}$ that the measure on vectors in the resolution

$$(\Psi', \Psi) = \int_{\mathfrak{S}} (\Psi', \Phi)\, d\mu\,(\Phi)(\Phi, \Psi) \quad (A1)$$

may be chosen invariant under all unitary transformations leaving $\mathfrak{S}$ unchanged. Let U be such a transformation:

$$U\mathfrak{S} = \mathfrak{S}. \quad (A2)$$

Then unitary invariance of (A1) combined with (A2) yields

$$(\Psi', \Psi) = \int_{\mathfrak{S}} (\Psi', U\Phi)\, d\mu\,(\Phi)(U\Phi, \Psi)$$

$$= \int_{\mathfrak{S}} (\Psi', \Phi)\, d\mu\,(U^{-1}\Phi)(\Phi, \Psi). \quad (A3)$$

Namely, if $d\mu(\Phi)$ was a suitable weighting, we see that an equally good weighting on vectors Φ is $d\mu(U^{-1}\Phi)$. If $d\mu(U^{-1}\Phi) = d\mu(\Phi)$, then the measure already possesses the desired invariance. If μ is not invariant we may proceed as follows:

It follows from (A2) that $U^p\mathfrak{S} = \mathfrak{S}$, where p is an arbitrary integer. The generalization of (A3) would imply that $d\mu(U^p\Phi)$ is as good a weighting

[3] V. Bargmann, Commun. Pure Appl. Math. **14**, 187 (1961); Proc. Natl. Acad. Sci. U. S. **48**, 199 (1962). For the closely related formalism, see reference 1, Sec. 2.C.

[4] One type of labeling for boson fields is discussed by H. Araki, J. Math. Phys. **1**, 492 (1960).

[5] J. R. Klauder, Ann. Phys. (NY) **11**, 123 (1960), especially p. 127, and pp. 142–153; S. S. Schweber, J. Math. Phys. **3**, 831 (1962).

[6] Such an approach to dynamics was discussed by J. von Neumann, "The Theory of the Positron," Lecture Notes, Institute for Advanced Study, Princeton, New Jersey, 1936.

[7] See reference 1, Sec. 4.

as $d\mu(\Phi)$. If U is a cyclic element of order P, i.e., $U^P = 1$, then the combination

$$d\mu_P(\Phi) \equiv \frac{1}{P}\sum_{p=0}^{P-1} d\mu\,(U^p\Phi) \tag{A4}$$

is invariant under U, $d\mu_P(U^{-1}\Phi) = d\mu_P(\Phi)$. Henceforth, we would use only the invariant form and rename the measure $d\mu(\Phi)$. This procedure can be extended to include invariance under all cyclic elements that leave $\mathfrak{S}$ invariant.

On the other hand, if U is not a cyclic element, then we shall regard it as an element of a one-parameter Lie group whose elements $U[\alpha]$ satisfy $U[\beta]U[\alpha] = U[\beta + \alpha]$. Invariance of the compact set $\mathfrak{S}$ under all powers of U leads by continuity and Postulate 1 to invariance of $\mathfrak{S}$ under the entire one-parameter subgroup containing U. We now can form a quantity analogous to (A4), namely,

$$d\mu_c(\Phi) = \frac{1}{\int d\alpha}\int d\mu\,(U[\alpha]\Phi)\,d\alpha, \tag{A5}$$

an expression which exists in virtue of the finite parameter range in compact groups. Clearly, $d\mu_c$ is invariant under $U[\beta]$, i.e., $d\mu_c(U[\beta]\Phi) = d\mu_c(\Phi)$. Again we rename this invariant form $d\mu(\Phi)$. By extending the preceding techniques to all invariant transformations of $\mathfrak{S}$, we establish Theorem 1.

While the above proof holds only for compact spaces, the left-invariant group measure as suggested by Theorem 1 may always be examined for individual noncompact spaces to see whether or not it satisfies (A1).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 8 AUGUST 1963

Continuous-Representation Theory. II. Generalized Relation between Quantum and Classical Dynamics

JOHN R. KLAUDER
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received 26 December 1962)

This paper discusses an application to the study of dynamics of the typical overcomplete, nonindependent sets of unit vectors that characterize continuous-representation theory. It is shown in particular that the conventional, classical Hamiltonian dynamical formalism arises from an analysis of quantum dynamics restricted to an overcomplete, nonindependent set of vectors which lie in one-to-one correspondence with, and are labeled by, points in phase space. A generalized "classical" mechanics is then defined by the extremal of the quantum-mechanical action functional with respect to a restricted set of unit vectors whose c-number labels become the dynamical variables. This kind of "classical" formalism is discussed in some generality, and is applied not only to simple single-particle problems, but also to finite-spin degrees of freedom and to fermion field oscillators. These latter cases are examples of an important class of problems called exact, for which a study of the classical dynamics alone is sufficient to infer the correct quantum dynamics.

1. INTRODUCTION

THE general postulates of continuous representations of Hilbert space have been stated elsewhere.[1] The overcomplete family of states, hereafter abbreviated to OFS and denoted by $\mathfrak{S}$, that is involved in such a representation may be visualized as forming one or more closed, connected "patches" on the unit sphere in Hilbert space defined by $||\Phi|| = 1$. In order to discuss dynamics and time evolution we shall define a *path* to be a continuous, unit-vector-valued time function $\Phi(t)$. Now, a general variation of the path $\Phi(t)$, apart from simple ray rotations (e.g., $\Phi(t) = \exp[i\lambda(t)]\Phi$), in the action functional

$$I = \int [i\hbar(\Phi, d\Phi/dt) - (\Phi, \mathcal{H}\Phi)]\,dt, \tag{1}$$

yields as the Euler–Lagrange equations, the Schrödinger equation of motion. However, we may ask what are the dynamical consequences if Eq. (1) is extremized over only a *restricted* set of paths, such as those constrained to lie in $\mathfrak{S}$: $\Phi(t) \in \mathfrak{S}$?

Through the study of a one-dimensional, single-particle problem in Sec. 2,[2] we conclude that when

[1] J. R. Klauder, J. Math. Phys. 4, 1055 (1963) (preceding paper) hereafter referred to as I.

[2] A brief discussion based on this example appears in J. R. Klauder, Helv. Phys. Acta 35, 333 (1962).

as $d\mu(\Phi)$. If U is a cyclic element of order P, i.e., $U^P = 1$, then the combination

$$d\mu_P(\Phi) \equiv \frac{1}{P}\sum_{p=0}^{P-1} d\mu\,(U^p\Phi) \tag{A4}$$

is invariant under U, $d\mu_P(U^{-1}\Phi) = d\mu_P(\Phi)$. Henceforth, we would use only the invariant form and rename the measure $d\mu(\Phi)$. This procedure can be extended to include invariance under all cyclic elements that leave $\mathfrak{S}$ invariant.

On the other hand, if U is not a cyclic element, then we shall regard it as an element of a one-parameter Lie group whose elements $U[\alpha]$ satisfy $U[\beta]U[\alpha] = U[\beta + \alpha]$. Invariance of the compact set $\mathfrak{S}$ under all powers of U leads by continuity and Postulate 1 to invariance of $\mathfrak{S}$ under the entire one-parameter subgroup containing U. We now can form a quantity analogous to (A4), namely,

$$d\mu_c(\Phi) = \frac{1}{\int d\alpha}\int d\mu\,(U[\alpha]\Phi)\,d\alpha, \tag{A5}$$

an expression which exists in virtue of the finite parameter range in compact groups. Clearly, $d\mu_c$ is invariant under $U[\beta]$, i.e., $d\mu_c(U[\beta]\Phi) = d\mu_c(\Phi)$. Again we rename this invariant form $d\mu(\Phi)$. By extending the preceding techniques to all invariant transformations of $\mathfrak{S}$, we establish Theorem 1.

While the above proof holds only for compact spaces, the left-invariant group measure as suggested by Theorem 1 may always be examined for individual noncompact spaces to see whether or not it satisfies (A1).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 8 AUGUST 1963

Continuous-Representation Theory. II. Generalized Relation between Quantum and Classical Dynamics

JOHN R. KLAUDER
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received 26 December 1962)

This paper discusses an application to the study of dynamics of the typical overcomplete, nonindependent sets of unit vectors that characterize continuous-representation theory. It is shown in particular that the conventional, classical Hamiltonian dynamical formalism arises from an analysis of quantum dynamics restricted to an overcomplete, nonindependent set of vectors which lie in one-to-one correspondence with, and are labeled by, points in phase space. A generalized "classical" mechanics is then defined by the extremal of the quantum-mechanical action functional with respect to a restricted set of unit vectors whose c-number labels become the dynamical variables. This kind of "classical" formalism is discussed in some generality, and is applied not only to simple single-particle problems, but also to finite-spin degrees of freedom and to fermion field oscillators. These latter cases are examples of an important class of problems called exact, for which a study of the classical dynamics alone is sufficient to infer the correct quantum dynamics.

1. INTRODUCTION

THE general postulates of continuous representations of Hilbert space have been stated elsewhere.[1] The overcomplete family of states, hereafter abbreviated to OFS and denoted by $\mathfrak{S}$, that is involved in such a representation may be visualized as forming one or more closed, connected "patches" on the unit sphere in Hilbert space defined by $||\Phi|| = 1$. In order to discuss dynamics and time evolution we shall define a *path* to be a continuous, unit-vector-valued time function $\Phi(t)$. Now, a general variation of the path $\Phi(t)$, apart from simple ray rotations (e.g., $\Phi(t) = \exp[i\lambda(t)]\Phi$), in the action functional

$$I = \int [i\hbar(\Phi, d\Phi/dt) - (\Phi, \mathcal{H}\Phi)]\,dt, \tag{1}$$

yields as the Euler–Lagrange equations, the Schrödinger equation of motion. However, we may ask what are the dynamical consequences if Eq. (1) is extremized over only a *restricted* set of paths, such as those constrained to lie in $\mathfrak{S}$: $\Phi(t) \in \mathfrak{S}$?

Through the study of a one-dimensional, single-particle problem in Sec. 2,[2] we conclude that when

[1] J. R. Klauder, J. Math. Phys. 4, 1055 (1963) (preceding paper) hereafter referred to as I.

[2] A brief discussion based on this example appears in J. R. Klauder, Helv. Phys. Acta 35, 333 (1962).

evaluated for a particular restricted set of unit vectors, parameterized or labeled by phase-space points p, q [the vectors of this set being given by (3)], the quantum action functional, Eq. (1), reduces in form to the *classical* action functional. That is, Eq. (1) reduces to

$$I = \int [p\dot{q} - H(p, q)]\, dt. \tag{2}$$

Consequently, by extremizing (2) with respect to just this special set of states, we would obtain the classical equations and not the quantum equations. While the action functional (2) and its extremal equations thus have the classical form, these dynamical variables are still to be interpreted as labels for Hilbert-space vectors. Likewise, the physical interpretation of the theory still follows the stochastic quantum prescription, i.e., vector inner products are transition amplitudes. Thus, the distinction between classical mechanics based on (2), and (an approximate form of) quantum mechanics also based on (2), lies in the *interpretation* of p and q and in how results are read out of the formalism; the quantum interpretation generally leads to an approximate form of quantum mechanics since (1) is described by (2) for only a *restricted* set of paths. We shall refer to an action functional of the form (2)—wherein the variables p and q are ordinary c-number time functions, but which are interpreted as labels for Hilbert-space vectors—as a "classical" action functional and the Euler–Lagrange equations derived therefrom as "classical" equations of motion. We observe, therefore, that merely by reinterpreting the c-number variables of the classical theory, we can view the classical action functional as a restricted evaluation of the true quantum mechanical action.

The above example, suggested by the study of a particular restricted set $\mathfrak{S}$, permits an obvious abstract extension to an arbitrary set $\mathfrak{S}$. Namely, *the "classical" equations of motion relative to* $\mathfrak{S}$ *arise as a result of extremizing* (1) *over just those vector functions for which* $\Phi(t) \in \mathfrak{S}$. Here we have a *relative* definition of the attribute "classical"—its relative nature depending on the size of the set $\mathfrak{S}$—that is generally applicable to any system with arbitrary statistics.

As an extreme situation, suppose $\mathfrak{S}$ were so large as to equal $\mathfrak{T}$, the set of all unit vectors. In that case, the resultant "classical" equations would be physically equivalent to the usual quantum-mechanical equations. It must be stressed, however, that it is not always necessary that the set $\mathfrak{S}$ be as large as $\mathfrak{T}$ in order that the "classical" equations be physically equivalent to the quantum equations. For example, if $\mathcal{H} \equiv 0$, then extremizing (1) over any (complete) set $\mathfrak{S}$ would lead to $\Phi(t) = \Phi(0)$ for each vector $\Phi(0) \in \mathfrak{S}$; the completeness of $\mathfrak{S}$ would then correctly imply the simple evolutionary behavior for any state vector in the Hilbert space $\mathfrak{H}$. This trivial example has the property that the *exact* solution to the quantum-mechanical equations, $\Phi(t)$, obeys $\Phi(t) \in \mathfrak{S}$ if only $\Phi(0) \in \mathfrak{S}$; namely, that the true extremal to (1) is a path that remains in $\mathfrak{S}$ if only it started in $\mathfrak{S}$. This possibility is by no means confined to the case $\mathcal{H} = 0$, and for the general case we introduce the

Definition: An *exact* "classical" action functional $I_{\mathfrak{S}}\{\Phi(t)\}$ relative to the set $\mathfrak{S}$ is one whose extremal solutions correctly correspond to true extremal vector-valued time functions. In other words, if the true quantum solution $\Phi(t) = e^{-i t \mathcal{H}/\hbar}\Phi(0)$ lies within the complete set $\mathfrak{S}$, $\Phi(t) \in \mathfrak{S}$, assuming only that it initially lay within $\mathfrak{S}$, then the "classical" theory is exact. If a "classical" action functional is not exact, then we shall call it inexact.

Clearly the existence of an exact "classical" action functional depends strongly on both the set $\mathfrak{S}$ and the Hamiltonian operator $\mathcal{H}$.

It follows as a corollary to the preceding definition that the *value* of the action evaluated for an extremal path will vanish for an exact "classical" action principle. Therefore, for an action functional to be exact, it is necessary but not sufficient (this latter aspect is further discussed in Sec. 2) that

$$I_{\mathfrak{S}}\{\Phi_{\text{extremal}}(t)\} = 0, \tag{E}$$

when $I_{\mathfrak{S}}$ is extremized with respect to all $\Phi(t) \in \mathfrak{S}$. We shall refer to this vanishing of I for exact action principles as "criterion E." If criterion E is not obeyed then, of course, the action principle is inexact.

It is the subject of this paper to study some properties of our generalized "classical" formalism and learn some of the consequences that arise from restricted variations of quantum-mechanical action functionals. In Sec. 2 we study in some detail the properties of a single-particle, nonrelativistic, one-dimensional example. For Hamiltonians linear in the momentum and position operators, an exact "classical" action functional arises. Oscillator Hamiltonian operators can also lead to exact cases but only for a unique choice of the set $\mathfrak{S}$. Other Hamiltonians lead to inexact equations of motion (unless $\mathfrak{S}$ becomes significantly enlarged). Our study of canonical transformations suggests that partial

physical significance of the dynamical variables is contained in the important *canonical kinematical form*, $i\hbar(\Phi, d\Phi)$, expressed as a function of the labels. A resolution of unity and the associated continuous representation is also discussed.

Generalization of the preceding analysis to an abstract N-dimensional Lie group of unitary transformations is the subject of Sec. 3. Various formulas are given for exact and inexact "classical" action functionals to discuss the "classical" formalism under certain special circumstances. Of special note, a fact which is stressed here possibly for the first time, is that classical theories with nonunique solutions may nevertheless have distinguished solutions that exhibit the exact character of their equations of motion, and therefore a study of these solutions may in turn shed light on the quantum analysis of classical theories with symmetries or with gauge freedoms. As an example, the "classical" action for a two-component-spin degree of freedom is presented.

In Sec. 4 we discuss a particularly simple labeling for a finite-dimensional Hilbert space where the labels are simply the vector components themselves. Application is made to a single-fermion oscillator, and generalization to a fermion field shows that Dirac action functionals for c-number spinor fields may be considered exact even in the presence of external sources.

Our viewpoint of "classical" mechanics is simply as quantum mechanics, evaluated for a restricted class of paths. Quantization is, therefore, already accomplished in part merely by the reinterpretation of classical variables as vector labels. The continuity of the labels and the corresponding continuity of the associated vectors plays a fundamental role in our dynamical viewpoint. It leads, for example, to an essential difference between our formalism and that of Schwinger[3] who considers only orthogonal vector or operator sets, and thus must exclude finite-dimensional Hilbert spaces. Other recent abstract dynamical studies include those of Sudarshan and coworkers,[4] which focus almost exclusively on operators forming an orthogonal operator basis and their time evolution, generalizing earlier work of Moyal.[5]

2. AN ELEMENTARY EXAMPLE

A. An OFS Parameterized by Phase Space Points

Let us consider a single, nonrelativistic particle free to move in only one dimension. We denote Hermitian position and momentum operators by Q and P, respectively; these operators satisfy

$$[Q, P] = i\hbar.$$

Along with these operators, let us introduce two c numbers q and p with the dimensions of coordinate and momentum, respectively. We now build the unitary operator

$$U[p, q] \equiv \exp(-iqP/\hbar) \exp(ipQ/\hbar). \tag{3a}$$

With the help of a fiducial unit vector Φ_0, we define

$$\Phi[p, q] \equiv U[p, q]\Phi_0; \tag{3b}$$

each $\Phi[p, q]$ is a unit vector and the set of these vectors for all p and q define $\mathfrak{S}$. Thus, in the present example, the label space consists of all points in phase space, and the mapping from label points to Hilbert-space vectors $p, q \rightarrow \Phi[p, q]$, is a homeomorphism explicitly displayed in Eqs. (3a) and (3b). We wish to emphasize that $\Phi[p, q]$ is not a particular *representation* of a Hilbert-space vector, but it is a vector in its own right.

In order for the set of states $\mathfrak{S}$ to be an OFS suitable for a continuous representation, it is necessary that $\mathfrak{S}$ and the labeling of the vectors therein satisfy three postulates.[1] The verification of Posttulate 1 of I regarding local density and continuity of the vectors in $\mathfrak{S}$, as well as Postulate 2 of I regarding labeling continuity may be established by a study of the quantity

$$\mathcal{K}(p', q'; p, q) \equiv (\Phi[p', q'], \Phi[p, q])$$
$$= e^{-ip'(q-q')/\hbar}(\Phi_0, e^{-i(q-q')P/\hbar}e^{i(p-p')Q/\hbar}\Phi_0). \tag{4}$$

$\mathcal{K}$ may be brought arbitrarily close to the value one for fixed p and q, and $\mathcal{K}$ is in addition separately continuous in p and q. Part of Postulate 3 of I, the completeness of the set $\mathfrak{S}$, has been shown by Moyal.[5] Implicit in his work is the fulfillment of the remaining condition of Postulate 3 of I, which will be discussed in part C below. Hence, the vectors, defined in Eq. (3) form an OFS. Rigorous proof of this fact will be given in Part IV in collaboration with J. KcKenna.

The preceding remarks are valid for any choice of Φ_0, independent of the fact that the set $\mathfrak{S}$ clearly depends on Φ_0. We wish now to eliminate some of the arbitrariness of Φ_0 so as to simplify the inter-

---

[3] J. Schwinger, Proc. Natl. Acad. Sci. U. S. **46**, 883, 1401 (1960).

[4] E. C. G. Sudarshan, *Brandeis Summer Institute Lecture Notes*," (W. A. Benjamin, Inc., New York, 1961), Vol. 2, p. 143. T. F. Jordan and E. C. G. Sudarshan, Rev. Mod. Phys. **33**, 515 (1961). Related ideas appear in: E. H. Wichmann, J. Math. Phys. **2**, 876 (1961); F. Bopp, *Heisenberg-Festschrift* (Frederick Vieweg und Sohn, Braunschweig, Germany, 1961), p. 128.

[5] J. E. Moyal, Proc. Cambridge Phil. Soc. **45**, 99 (1949).

pretation of p and q. In particular, we ask that

$$(\Phi[p, q], P\Phi[p, q]) = p, \tag{5a}$$

$$(\Phi[p, q], Q\Phi[p, q]) = q, \tag{5b}$$

two relations which impose on Φ_0 the modest restrictions

$$(\Phi_0, P\Phi_0) = 0, \tag{6a}$$

$$(\Phi_0, Q\Phi_0) = 0. \tag{6b}$$

Equation (6) follows as a consequence of (5) directly, with the help of the familiar equality

$$U^{-1}[p, q](\alpha Q + \beta P)U[p, q] = \alpha(Q + q) + \beta(P + p),$$

where α and β are arbitrary c numbers. Henceforth, we shall assume Eqs. (5) and (6) to be satisfied.

The adoption of (5) and (6) narrows the possible forms that can be taken by $\mathcal{K}$ in Eq. (4). In particular, Eqs. (5) and (6) lead to the following canonical kinematical form:

$$i\hbar(\Phi[p, q], d\Phi[p, q]) = p\, dq, \tag{7}$$

for a kind of differential form for $\mathcal{K}$. We now take up several aspects of this example in more detail.

Relation of Quantum and "Classical" Dynamics[2]

In order to discuss restricted dynamical equations with the aid of the action functional (1), we let the unit vectors $\Phi[p, q] \in \mathfrak{S}$ be functions of time. This we do simply by permitting p and q to be arbitrary, independent c-number time functions $p(t)$ and $q(t)$, and define

$$\Phi(t) \equiv \Phi[p(t), q(t)]. \tag{8}$$

Note that the operators P and Q remain unchanged here. Combining (7) and (8) we find

$$i\hbar(\Phi, d\Phi/dt) = p\dot{q};$$

here, and elsewhere, the dot signifies a time derivative, $\dot{q} \equiv dq/dt$.

With the help of the preceding expression, Eq. (1) reduces to

$$I_{\mathfrak{S}} = \int [p\dot{q} - H(p, q)]\, dt, \tag{9}$$

where

$$\begin{aligned} H(p, q) &\equiv (\Phi[p, q], \mathcal{H}(P, Q)\Phi[p, q]) \\ &= (\Phi_0, \mathcal{H}(P + p, Q + q)\Phi_0) \\ &\equiv \mathcal{H}(p, q) + \mathcal{O}(\hbar; \Phi_0; p, q). \end{aligned} \tag{10}$$

Equation (9) has the form of a *classical* action functional where the classical Hamiltonian is $H(p, q)$. According to (10), we see that $H(p, q)$ has the functional form of the quantum mechanical Hamiltonian with explicit c-number substitution, i.e., $\mathcal{H}(p, q)$ plus an additional term $\mathcal{O}$ depending on $\hbar$, the fiducial vector Φ_0 as well as on the momentum and coordinate. For nonpathological Hamiltonian operators, $\mathcal{O}$ depends only on positive powers of $\hbar$; hence in this case,

$$\lim_{\hbar \to 0} \mathcal{O}(\hbar; \Phi_0; p, q) = 0.$$

Thus as $\hbar \to 0$ we obtain $H(p, q) = \mathcal{H}(p, q)$, which is just the conventional relation in order that $H(p, q)$ be the appropriate classical Hamiltonian for the system under discussion. In this same limit, p and q achieve their conventional, classical sharp physical significance since, e.g.,

$$\lim_{\hbar \to 0} (\Phi[p, q], (P - p)^2\Phi[p, q]) = 0.$$

When $\hbar \to 0$, it is clear that a stationary variation of (9) yields the conventional classical equations of motion, and not the quantum equations. For a macroscopic system, where $\mathcal{H}(p, q) \gg \mathcal{O}$, we expect the classical equations and interpretation to be very accurate. Hence we have established that a restricted variation of I over just those vectors $\Phi(t)$ which obey (8) yields essentially the classical equations of motion, for macroscopic systems, and in the limit $\hbar \to 0$, it yields precisely the entire classical "picture."

But such limiting cases are not the only ones in which the conventional classical equations can arise. We shall shortly prove that, for any Hamiltonian operator of the form $\mathcal{H} = \frac{1}{2}P^2 + V(Q)$, we can always choose Φ_0, consistent with (6), so that $\mathcal{O}$ *is as small as desired* for any system and not only for macroscopic systems. Thus, with $\mathcal{O}$ small, we again recover the conventional classical equations of motion, but since $\hbar \neq 0$, p and q do not have a sharp physical meaning, and these "classical" variables must be correctly interpreted for what they are: labels for Hilbert-space vectors as in (3). Thus, simply by a reinterpretation of the classical variables, we can regard the classical action functional as a restricted evaluation of the quantum action. With this wider viewpoint understood, $\mathcal{O}$ ceases to be conceptually bothersome, and we can just as well regard $H(p, q)$ itself as the "classical" Hamiltonian; indeed, the harmonic oscillator is a very important case where $\mathcal{O}$ is best chosen not to vanish.

Certain classical statements contained in the present formalism may be compared with those

predicted by the conventional viewpoint based on Ehrenfest's theorem. This theorem states[6] that the expectation values

$$\bar{P} \equiv (\Phi(t), P\Phi(t)), \tag{11a}$$

$$\bar{Q} \equiv (\Phi(t), Q\Phi(t)), \tag{11b}$$

satisfy Hamilton's equations when $\Phi(t)$ is the true quantum solution. Clearly, for exact action principles when the true solution remains within $\mathfrak{S}$ the quantities defined in (11) are just those defined in (5). For inexact action principles, however, (5) differs from (11). From the point of view of this paper, the equations of motion that follow from (9) are of precisely the conventional form, e.g.,

$$dp/dt = -\partial H(p, q)/\partial q.$$

On the other hand, Ehrenfest's theorem leads to

$$d\bar{P}/dt = i(\Phi(t), [\mathfrak{H}, P]\Phi(t)) = \overline{i[\mathfrak{H}, P]} \neq -\partial\mathfrak{H}(\bar{P}, \bar{Q})/\partial\bar{Q},$$

the inequality holding as a general statement. Therefore, while the label point of view for position and momentum leads to Hamilton's equations, the expectation point of view in (11) for such variables leads to Hamilton's equations only in the limit $\hbar \to 0$.

Additional information regarding the "classical" dynamics of our one-dimensional examples may more easily be found if we first modify the set $\mathfrak{S}$ in a very simple manner.

B. Modification of the OFS to Include Phase Factors

Instead of the set of vectors defined in Eq. (3), let us choose

$$\Phi[p, q, \alpha] \equiv U[p, q, \alpha]\Phi_0 \equiv \exp(-i\alpha/\hbar) \times \exp(-iqP/\hbar) \exp(ipQ/\hbar)\Phi_0 \tag{12}$$

to be members of $\mathfrak{S}$, where α, $0 \leq \alpha < 2\pi\hbar$ is a new c-number label. This expanded set of vectors may be treated exactly as in part A above. In particular, if p, q, and α become functions of time, the action functional for this restricted set of unit vectors has the form

$$I_{\mathfrak{S}} = \int [p\dot{q} + \dot{\alpha} - H(p, q)]\, dt, \tag{13}$$

where $H(p, q)$ is the same as in (10). Clearly α in no way alters the dynamics of p and q, nor is the evolution of α determined by extremizing the action (13). If $\alpha(t)$ remains a free and undetermined function, then Eq. (13) evaluated for "an extremal path" has not one but many values, which is why arbitrary ray rotations were excluded in the consideration of Eq. (1). Instead of regarding α as an independent "classical" dynamical variable having undetermined behavior, we shall elaborate a different viewpoint.

Along with the unrestricted set $\mathfrak{S}$ defined by Eq. (12), let us also consider a family of its subsets restricted so that $\alpha = \alpha(p, q, t)$ for various functions α. Each of these restricted sets remains a valid OFS; e.g., if we put $\alpha = 0$, we recover the set $\mathfrak{S}$ discussed in part A. We now consider several applications of various OFS restricted by $\alpha = \alpha(p, q, t)$ where the form of α is chosen to accomplish one or another specific purpose, e.g., to share with the restricted set a certain desirable property exhibited in the unrestricted set.

Exact "Classical" Action Functions

We wish to demonstrate that for linear and quadratic Hamiltonians it is possible to choose $\alpha = \alpha(p, q)$ in (12) so as to obtain a determinate, exact "classical" system. This is possible in cases where the unrestricted set (12) contains the evolution of its own members.

As a first example let us choose $\mathfrak{H} = AP$, where A is a constant. It is clear that

$$\Phi(t) = e^{-itAP/\hbar}\Phi(0) = \Phi[p(0), q(0) + tA, \alpha(0)],$$

so that $\Phi(t) \in \mathfrak{S}$. The phase α does not change at all here so we may set $\alpha = 0$ for this example, picking out the OFS discussed in A. Since $H(p, q) = Ap$, it is clear that criterion E is satisfied.

As a second example, let $\mathfrak{H} = BQ$, where B is a constant. Then

$$\Phi(t) = e^{-itBQ/\hbar}\Phi(0) = \Phi[p(0) - tB, q(0), \alpha(0) + tBq(0)], \tag{14}$$

a vector which is of the general form of (12) for all t. If we evaluate (13) for the indicated functions p, q, and α, we see that criterion E is obeyed. An alternate way to express $\alpha(t)$ is $\alpha = -pq$, where we arbitrarily choose $\alpha(0) = -p(0)q(0)$. Thus, the set of vectors of the form (12) restricted so that $\alpha = -pq$ also contains the evolution of its members for the Hamiltonian $\mathfrak{H} = BQ$. The "classical" action assumes the form $-\int q(\dot{p} + B)\, dt$ which clearly satisfies criterion E. Other choices of $\alpha = \alpha(p, q, t)$ are appropriate for the general linear case $\mathfrak{H}_{\text{lin}} =$

[6] See, e.g., L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1955), 2nd ed., p. 25.

$AP + BQ + C$, to exhibit the exact character of $I_{\mathfrak{S}}$. For any linear Hamiltonian, $\mathcal{O}$ in Eq. (10) vanishes.

The fact that (13) may be exact for linear Hamiltonians can be seen as follows. The unitary transformations $U[p, q, \alpha]$ acting on Φ_0 in (12) form a group in virtue of the closed Lie algebra of 1, P, and Q; specifically,

$$U[p, q, \alpha]U[p', q', \alpha'] = U[p + p', q + q', \alpha + \alpha' - pq'], \quad (15a)$$

which, on introducing Φ_0, states that

$$U[p, q, \alpha]\Phi[p', q', \alpha'] = \Phi[p + p', q + q', \alpha + \alpha' - pq']. \quad (15b)$$

Hence, for any evolution operator $\exp(-it\mathcal{H}/\hbar)$ that can be expressed in the form $U[p, q, \alpha]$, a choice for $\alpha(t)$ [that implicit in (15b)] exists so that the associated OFS leads to an exact action principle satisfying criterion E. Such $\mathcal{H}$ are of the linear form $AP + BQ + C$. A generalization to coefficients A, B, and C that are functions of time offers no difficulty.

To illustrate the analysis of quadratic Hamiltonians, we consider only the particular quadratic form $\mathcal{H}_{\text{h.o.}} = \frac{1}{2}(P^2 + \omega^2 Q^2 - \hbar\omega)$ and we choose $\alpha = -\frac{1}{2}pq$ in (12). [When linear driving terms are present other choices for $\alpha(t)$ are needed to exhibit the exact nature of a subset of the unrestricted set $\mathfrak{S}$.] With this choice for α we are effectively considering the set $\mathfrak{S}$ whose vectors are

$$\Phi[p, q] = \exp[-i(qP - pQ)/\hbar]\Phi_0, \quad (16)$$

and the corresponding classical action functional

$$I_{\mathfrak{S}} = \int [\tfrac{1}{2}(p\dot{q} - q\dot{p}) - H(p, q)]\, dt. \quad (17)$$

Consider now the exactness of this action principle. The quantum evolution of the states in (16) is

$$\Phi(t) = e^{-it\mathcal{H}_{\text{h.o.}}/\hbar}\Phi(0) = \exp\{-i[q(t)P - p(t)Q]/\hbar\}e^{-it\mathcal{H}_{\text{h.o.}}/\hbar}\Phi_0, \quad (18)$$

where

$$q(t) = q(0)\cos\omega t + \omega^{-1}p(0)\sin\omega t,$$
$$p(t) = -\omega q(0)\sin\omega t + p(0)\cos\omega t.$$

It follows from (18) that $\Phi(t) \in \mathfrak{S}$, i.e., it is of the form (16), if and only if Φ_0 is the ground state of $\mathcal{H}_{\text{h.o.}}$, although other eigenstates of $\mathcal{H}_{\text{h.o.}}$ would only give phase factors that could easily be absorbed into $\alpha(t)$. By choosing Φ_0 as an eigenstate, the addition of $\mathcal{H}_{\text{h.o.}}$ to the Lie algebra of 1, P, and Q can thus be *effected* without its actual *inclusion*, again giving rise to an exact "classical" action functional.

Armed with this result, we now observe the interesting conclusion for oscillator Hamiltonians that an exact "classical" theory arises by extremizing the "classical" action function's dependence on the *fiducial vector* Φ_0 and thus ensuring the eigenstate property of Φ_0; a global extremization further narrows Φ_0 to be the ground state of the oscillator. With this choice, $\mathcal{O} = \frac{1}{2}\hbar\omega$. The "classical" Hamilton Eq. (10) then becomes $H_{\text{h.o.}} = \frac{1}{2}(p^2 + \omega^2 q^2)$ corresponding to the preceding form for the harmonic-oscillator Hamiltonian, $\mathcal{H}_{\text{h.o.}}$, in which the zero-point energy was subtracted off (a physically very attractive correspondence indeed when generalized to a Bose field!). Thus, (i) the classical equations for a harmonic oscillator imply the correct quantum-mechanical time automorphism, (ii) the spectrum of the corresponding quantum Hamiltonian begins at zero, without a zero-point subtraction being necessary, and (iii) criterion E is evidently satisfied in (17) by virtue of Euler's theorem on homogeneous forms.

The success of the analysis in (18) mainly lay in being able to commute the evolutionary operator with the unitary operator in (16) with the only "cost" being a change of the labels. This situation will prevail if the evolutionary operator generates a family of outer automorphisms of the Lie algebra of 1, P, and Q, which is only true for a general quadratic Hamiltonian.[7]

Canonical Transformations and Inexact "Classical" Action Functionals

While criterion E is necessary to have an exact "classical" action functional it is not sufficient for by a suitable choice of $\alpha = \alpha(p, q)$ in (12) and in (13) we can always satisfy criterion E. In particular, we could choose α equal to $F(\mathfrak{q}, q)$, a function related to Hamilton's characteristic function, chosen such that $p \equiv -\partial F/\partial q$ and $\mathfrak{p} \equiv \partial F/\partial \mathfrak{q} = H(-\partial F/\partial q, q)$. After such a canonical transformation, the "classical" action function reads[8]

$$I_{\mathfrak{S}} = \int (\mathfrak{p}\dot{\mathfrak{q}} - \mathfrak{p})\, dt,$$

which clearly fulfills criterion E. The restricted subset of (12) determined by the above rule is an OFS conveniently labeled by $\mathfrak{p}$, $\mathfrak{q}$, $\Phi[\mathfrak{p}, \mathfrak{q}]$, in terms of

[7] J. E. Moyal and M. S. Bartlett, Proc. Cambridge Phil. Soc. 45, 545 (1949).

[8] See, e.g., H. Goldstein, *Classical Mechanics* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1950), Chaps. 8 and 9.

which $i\hbar(\Phi, d\Phi) = \mathfrak{p}\, d\mathfrak{q}$ consistent with (7). But the functional form of Eq. (5)—let alone that of the more involved relation (4)—would in general be very different when expressed in terms of the variables $\mathfrak{p}$, $\mathfrak{q}$. This difference reflects the fact that Eqs. (4) and (5) give to the canonical pair p, q a certain physical significance that may not be shared by the canonical pair $\mathfrak{p}$, $\mathfrak{q}$. With this physical difference recognized, it is clear that the Hamiltonian $\mathfrak{p}$ would no longer be the expectation value of an infinitesimal element that generates an Abelian subgroup of the set $U[p, q, \alpha]$ as (5a) represents. Instead, $\mathfrak{p}$ is the expectation value of the Hamiltonian $\mathcal{H}(P, Q)$ in the state $\Phi[\mathfrak{p}, \mathfrak{q}]$. If the Hamiltonian were not one of the linear or quadratic ones we discussed above, then the action functional would be inexact even though criterion E were fulfilled. These results in no way prohibit $\mathfrak{p}$ and $\mathfrak{q}$ from being "good" labels; they simply call our attention to the fact that (7) has many solutions [including all sets arising from (12) by the restriction $\alpha = \alpha(p, q)$], and among these is the solution (3) for variables p and q having a physical, translational invariance as is implied by (4).

It is worth emphasizing the basic relation of a classical canonical transformation to an associated quantum transformation from our point of view. To carry out any canonical transformation, we merely pass from one OFS to another OFS differing trivially from the first by having different phase factors; no involved unitary transformation acting both on operators and vectors is coupled to the classical canonical transformation. Such unitary transformations are a separate invariance group of the quantum theory.

By our definition, an inexact "classical" action functional is one whose associated OFS does not contain the true quantum-dynamical evolution of its members. The example above involving $\mathfrak{p}$ and $\mathfrak{q}$ implies that merely changing the labels of the vectors and introducing phase factors cannot, in general, make an inexact action become exact. In order to conclude that a given "classical" action functional satisfying criterion E is exact or not requires some additional information regarding the physical significance of the variables in which it is expressed. One convenient way to analyze "classical" action functionals—and that which we follow in this and in subsequent sections of this paper—is, by means of Eq. (1), to express the action functionals directly in terms of *specific* labels whose physical significance is implicitly contained in inner products such as (4).

The discussion associated with Eq. (18) has shown us that an inexact "classical" theory will result when $\mathcal{H}$ is neither linear nor quadratic in P and Q. There is even no choice of Φ_0 that will lead to an exact theory. However, in the case

$$\mathcal{H} = \tfrac{1}{2}P^2 + V(Q),$$

a suggestive choice for Φ_0 can be put forward that gives to the "classical" theory its conventional form. For this class of Hamiltonians we find from (10) that

$$\mathcal{O} = \tfrac{1}{2}(\Phi_0, P^2\Phi_0) + (\Phi_0, [V(Q + q) - V(q)]\Phi_0) \equiv c + v(q).$$

Thus $q(t)$ is the only dynamical variable on which $\mathcal{O}$ depends. By choosing Φ_0 sharp in Q space about zero we can make $v(q)$ arbitrarily small. In addition to Eq. (6), such a Φ_0 satisfies the relation $(\Phi_0, Q^2\Phi_0) =$ arbitrarily small. The price for reducing $v(q)$ to a negligible quantity is that now $c \equiv \tfrac{1}{2}(\Phi_0, P^2\Phi_0)$ becomes arbitrarily large. But we can cancel this constant by the choice of phase $\alpha = ct$ in (13), thus eliminating $\mathcal{O}$ altogether. In summary, if we (i) choose Φ_0 arbitrarily sharp in Q space about zero, and (ii) use a set of states including $\Phi[p, q, ct]$ as defined in (12), then we can bring the "classical" Hamiltonian $H(p, q)$ arbitrarily close to the conventional form $\frac{1}{2}p^2 + V(q)$ even when $\hbar \neq 0$.

Operationally we can argue that the choice of Φ_0 to make $v(q)$ negligible is a result of extremizing the "classical" action functional with respect to Φ_0, as was the case for the harmonic oscillator. In the present case we simply give priority to those parts of $\mathcal{O}$ that do not lead to surface terms.

A further investigation of inexact "classical" action functionals will be the subject of a separate study. There we shall consider the relative accuracy of the approximate vector solution $\Phi[p(t), q(t)]$, where p and q are solutions of the extremal equations based on (13), as compared to the true quantum-mechanical solution $\Phi(t) = \exp(-it\mathcal{H}/\hbar)\Phi(0)$. We anticipate that the approximate solutions will possess some form of "maximum accuracy" compared to the true solutions when we choose Φ_0 to extremize the "classical" action functional.

C. Resolution of Unity and Continuous Representations

While the phase variable α is itself eliminated in favor of some specific functional form $\alpha = \alpha(p, q, t)$,

in general the precise form of the elimination cannot be made *a priori* until the Hamiltonian is selected. It is fitting, therefore, that the resolution of unity in Eq. (1) of I is most directly found when all α are included. The completeness of the vectors in $\mathfrak{S}$ (α plays no role here, of course) has been demonstrated by Moyal, and is easily proved by taking recourse to a Schrödinger representation of Hilbert space. We wish rather to illustrate the utility of Theorem 1, which is discussed in I, in deriving the measure on vectors in $\mathfrak{S}$ in a resolution of unity.

Quite generally, the resolution of unity in terms of such vectors will have the form

$$1 = \int \Phi[p, q, \alpha]\Delta(p, q, \alpha)\, dp\, dq\, d\alpha\, \Phi^\dagger[p, q, \alpha]. \quad (19a)$$

Because of the group property in (15a) it is clear that the $U[p, q, \alpha]$ are also the elements of the unitary transformation group that leave $\mathfrak{S}$ invariant. It is further clear that they form a transitive permutation group acting in $\mathfrak{S}$ and in view of Theorem 1 of I, if Δ exists, it can be determined everywhere up to a constant directly from the left-invariant group measure.[9] A simple calculation shows that Δ is constant, i.e., the weighting in (19a) is independent of p, q, and α. That the weighting would be independent of α could have been anticipated since the elements $U[0, 0, \alpha]$ form an invariant subgroup of the set $U[p, q, \alpha]$ with an additive law of combination. Further, it is clear that α also disappears from the special integrand $\Phi\Phi^\dagger$ in (19a), so that the integral over α simply multiplies Δ by a factor. It is important that this scaling of Δ is by a *finite* factor, which follows from the periodic nature of the parameter α. Consequently, the resolution of unity assumes the form

$$1 = \int \Phi[p, q] \frac{dp\, dq}{2\pi\hbar} \Phi^\dagger[p, q], \quad (19b)$$

the over-all constant being determined, for example, by the single requirement that the expectation value of (19b) with Φ_0 is one. Although obtained and used in different ways, Eq. (19b) is a result which agrees with the solution in reference 5. Thus, while α has no fundamental dynamical role, it is extremely useful in deducing the resolution of unity (19b) expressed as an integration over the true dynamical variables. It is observed that the functional form of the resolution in (19b) is invariant under canonical transformations.[10]

The existence of (19b) as a valid resolution of unity shows that a representation of Hilbert space can be realized by a suitable class of phase space functions. In particular,

$$(\Psi_1, \Psi_2) = \int \psi_1^*(p, q) \frac{dp\, dq}{2\pi\hbar} \psi_2(p, q), \quad (20a)$$

where, according to (19b),

$$\psi(p, q) \equiv (\Phi[p, q], \Psi).$$

This definition for $\psi(p, q)$ does not lead to a vector space of arbitrary functions but rather to one composed of continuous functions that fulfill the relation

$$\psi(p', q') = (2\pi\hbar)^{-1} \int \mathcal{K}(p', q'; p, q)\psi(p, q)\, dp\, dq,$$

where $\mathcal{K}$ is defined in (4). This restriction on ψ is implied by (20a) when we set $\Psi_1 = \Phi[p', q']$. Among the properties required of ψ is the bound

$$|\psi(p, q)| \leq ||\Psi||,$$

since Φ is a unit vector.

The preceding representation reduces essentially to the Fock representation by entire analytic functions[11] when Φ_0 is chosen as the ground state of an harmonic oscillator. To see this, let us first reexpress the vectors in (3) in an equivalent form:

$$\Phi[p, q] = e^{-\frac{1}{2}(x^2+y^2)+ixy} e^{(x-iy)A^*} e^{-(x+iy)A} \Phi_0,$$

where

$$x \equiv (\tfrac{1}{2}\omega/\hbar)^{\frac{1}{2}} q, \qquad y \equiv -(\tfrac{1}{2}\omega/\hbar)^{\frac{1}{2}} p,$$
$$A \equiv (\tfrac{1}{2}\omega/\hbar)^{\frac{1}{2}} Q + i(\tfrac{1}{2}\omega\hbar)^{\frac{1}{2}} P,$$

and A^* is the adjoint of A. The operator A is the usual annihilation operator for an oscillator. If Φ_0 is the oscillator ground state, then $A\Phi_0 = 0$, and the related term in the expression for Φ will vanish. Apart from multiplicative normalization factors, therefore, $\Phi[p, q]$ depends only on $x - iy$.[12] To

[9] The invariant group measure is discussed, e.g., by E. P. Wigner, *Group Theory* (Academic Press Inc., New York, 1959), Chap. 10.

[10] It is worth remarking at this point why we do not consider a set $\mathfrak{S}$ that includes $\Phi[p] = \exp(ipQ)\Phi^0$ for all p. If this set $\mathfrak{S}$ is not complete then its utility is severely impaired; if it is considered complete then a resolution of unity in terms of these states should exist. Recourse to a Schrödinger representation shows that any kernel $K(x', x'')$ proposed as a matrix representation of unity fails to satisfy translational invariance unless Φ_0 is an eigenvector of the momentum operator P. Since such eigenvectors do not exist, one would be forced into resolutions of unity having physically undesirable characteristics.

[11] V. Bargmann, Commun. Pure Appl. Math. **14**, 187 (1961).

[12] A related set of normalized states, which depends essentially only on one complex variable, is discussed by J. R. Klauder, Ann. Phys. (NY) **11**, 123 (1960), p. 125.

eliminate the normalization scale factors, let us introduce

$$w \equiv (\Phi[p, q], \Phi_0);$$

in view of our choice for Φ_0, w never vanishes. Then

$$f(p, q) \equiv w^{-1}(\Phi[p, q], \Psi) = f(z),$$

i.e., f is a function of $z = x + iy$, or stated otherwise, f is analytic, which furthermore is defined everywhere. To account for the weight factor w we have introduced, we must redefine the inner product (20a) as

$$(\Psi_1, \Psi_2) = \int f_1^*(z)\, |w|^2 \frac{dx\, dy}{\pi} f_2(z), \qquad (20b)$$

where $|w|^2 = \exp[-(x^2 + y^2)]$. These results conform with those given by Bargmann.[11] In the representation associated with (20b), the functions $f(z)$ need only be entire; in a manner of speaking, the measure now contains the boundedness property required of ψ.

Whether or not we choose Φ_0 as above, we wish to emphasize that Eq. (20) entails a representation of *vectors* by phase-space functions as contrasted with the more conventional representation of *operators* by phase-space functions.[13] In the present formalism, operators are continuous functions of two phase-space points, e.g., for (20a)

$$\mathfrak{B}(p', q'; p, q) \equiv (\Phi[p', q'], \mathfrak{B}\Phi[p, q]),$$

and their operation on vectors is effected by integrating $\mathfrak{B}\psi$ over both p and q:

$$(\mathfrak{B}\psi)(p', q') = (2\pi\hbar)^{-1} \int \mathfrak{B}(p', q'; p, q)\, dp\, dq\, \psi(p, q)$$

[cf. Eqs. (7) and (8) of I]. Equation (9) of I ensures that the representation of $\mathfrak{B}$ is both continuous and unique.

3. LABELING BY PARAMETERS OF UNITARY LIE GROUPS

Let us consider a generalization of the examples discussed in the last section to the case of an N-dimensional Lie group of unitary transformations acting in an n-dimensional Hilbert space. An N-dimensional Lie group element is characterized by N parameters, l^a, $a = 1, 2, \cdots, N$. Elements near to unity may be generated from N skew-Hermitian infinitesimal elements L_a. The L_a are assumed to be elements of a Lie algebra whose commutator product satisfies the well-known conditions

$$[L_a, L_b] = c_{ab}^d L_d, \qquad (21)$$

wherein the summation convention for label indices has been adopted. For the sake of clarity and to facilitate the comparison with the preceding section, we shall assume our labels to be the so-called "canonical coordinates."[14] In terms of these labels, the finite unitary transformation

$$V[l^a] = \exp(l^a L_a).$$

The set of states $\mathfrak{S}$ is now defined to contain all vectors of the form

$$\Phi[l^a] = V[l^a]\Phi_0 = \exp(l^a L_a)\Phi_0. \qquad (22)$$

Postulates 1 and 2 of I may be verified by a study of the quantity

$$\mathfrak{K}(l'^a; l^a) \equiv (\Phi[l'^a], \Phi[l^a]), \qquad (23)$$

a continuous function of the single-parameter set $(l'^{-1} \cdot l)^a$, where the dot denotes group multiplication in label space. We remark on Postulate 3 of I below.

Considered as a function of time, we let

$$\Phi(t) = \Phi[l^a(t)] = \exp[l^a(t)L_a]\Phi_0.$$

By making use of the general operator rule

$$e^{(A+B)} = e^A + \int_0^1 e^{sA} B e^{(1-s)A}\, ds,$$

valid to first order in B, we find that

$$d\Phi = dl^b(t) M_b^c(t) L_c \Phi(t), \qquad (24)$$

where the numerical coefficients M_b^c are defined through the relation

$$M_b^c(t) L_c = \int_0^1 \exp[sl^a(t)L_a] L_b \exp[-sl^c(t)L_c]\, ds. \qquad (25)$$

Now let us introduce additional numerical coefficients U_c^d by

$$U_c^d(t) L_d = \exp[-l^a(t)L_a] L_c \exp[l^b(t)L_b]. \qquad (26a)$$

An implicit expression for the label space matrix $U = \{U_c^d\}$ is given by

$$U(t) = \exp[-l^a(t)c_a], \qquad (26b)$$

where c_a is the matrix formed from the structure constants whose bd element is c_{ab}^d. In terms of the coefficients U_c^d, we have

$$(\Phi(t), L_c\Phi(t)) = U_c^d(t)(\Phi_0, L_d\Phi_0) \equiv -iU_c^d(t)v_d/\hbar. \qquad (27)$$

The constants v_d are real and characterize the expectation value of $i\hbar L_d$ in the fiducial state Φ_0.

[13] E. P. Wigner, Phys. Rev. 40, 749 (1932); see also references 3–5, and additional references therein.

[14] C. Chevalley, *Theory of Lie Groups I* (Princeton University Press, Princeton, New Jersey, 1946), Chaps. IV and V.

Depending on the relative dimensions of N and n, the constants v_d may or may not determine Φ_0 up to a phase factor given an explicit representation of the infinitesimal elements. At any rate, in terms of the quantities defined above, the canonical kinematical form is

$$i\hbar(\Phi, d\Phi) \equiv y_b(l^a)\, dl^b, \tag{28}$$

where

$$y_b \equiv M^c_b U^d_c v_d. \tag{29a}$$

Adopting a matrix notation once again, we may express the "vector" y in terms of the "vector" v as

$$y = [1 - \exp(-l^a c_a)](l^a c_a)^{-1} v, \tag{29b}$$

where again c_a is the matrix of structure constants.[14]

The "classical" Hamiltonian is defined by

$$H(l^a) \equiv (\Phi[l^a], \mathcal{H}\Phi[l^a]). \tag{30a}$$

In the event that $\mathcal{H}$ is a linear sum of infinitesimal elements, i.e., $\mathcal{H} = i\hbar h^c L_c$, then from (27) we have

$$H(l^a) = h^c U^d_c(l^a) v_d. \tag{30b}$$

In either case, the "classical" action functional assumes the form

$$I_{\mathfrak{S}} = \int [y_b(l^a)\dot{l}^b - H(l^a)]\, dt. \tag{31}$$

The "classical" equations of motion follow from extremizing (31) with respect to independent variations in l^b. These equations are

$$A_{cb}\dot{l}^b \equiv \left(\frac{\partial y_b}{\partial l^c} - \frac{\partial y_c}{\partial l^b}\right)\dot{l}^b = \frac{\partial H}{\partial l^c}. \tag{32}$$

A more explicit form for A_{cb} may be found as follows. The time derivative of (26a) may be expressed with the aid of (25):

$$\begin{aligned}\dot{U}^e_c(t)L_e &= \dot{l}^b M^d_b \exp(-l^a L_a)[L_c, L_d] \exp(l^e L_e) \\ &= \dot{l}^b M^d_b c^f_{cd} U^e_f L_e.\end{aligned}$$

This relation is true for all time functions $l^b(t)$, and due to their linear independence and to the linear independence of the infinitesimal elements L_e, we have

$$\partial U^e_c/\partial l^b = M^d_b c^f_{cd} U^e_f. \tag{33}$$

In view of (29a), the partial derivative $\partial y_b/\partial l^c$ is given by

$$\frac{\partial y_b}{\partial l^c} = \frac{\partial M^f_b}{\partial l^c} U^e_f v_e + M^a_b \frac{\partial U^e_a}{\partial l^c} v_e.$$

Substituting from (33) and antisymmetrizing in c and b we finally obtain

$$\begin{aligned}A_{cb} = [&\partial M^f_b/\partial l^c - \partial M^f_c/\partial l^b \\ &+ M^a_b M^d_c c^f_{ad} - M^a_c M^d_b c^f_{ad}] U^e_f v_e.\end{aligned} \tag{34}$$

Returning to Eq. (32), we see from the antisymmetry of A_{cb} that

$$\dot{l}^c A_{cb} \dot{l}^b = \dot{l}^c \frac{\partial H}{\partial l^c} = \frac{dH}{dt} - \frac{\partial H}{\partial t} = 0,$$

which expresses the constancy of H if H is not an explicit function of t. If A_{cb} is nonsingular, Eq. (32) fully determines the solution. Conversely, if A_{cb} is singular, the equations of motion do not determine the solution $l^b(t)$ uniquely; furthermore, if the dimension N of the Lie group is odd, A is necessarily singular. Such was the case for the dynamical example involving p, q, and α in Sec. 2. But just as there was a distinguished choice for $\alpha(t)$ for linear Hamiltonians to make manifest the exact nature of their action functionals, we find an analogous distinguished solution for (32) whenever H has the form (30b), even if A_{cb} is singular. While an explicit form for this solution $l^a(t)$ is difficult to write down, it is clearly defined for any Φ_0 through the relation

$$\Phi(t) = \exp(th^b L_b) \exp[l^a(0)L_a]\Phi_0 \equiv \exp[l^a(t)L_a]\Phi_0. \tag{35}$$

That the time evolution in (35) remains a vector in $\mathfrak{S}$ is a consequence of the Baker–Hausdorff theorem, but closed-form solutions are available for only a few algebras.[15] The exact quantum-mechanical solution (35) is also the extremal solution for (31), and the evaluation of $I_{\mathfrak{S}}$ for this solution vanishes, thus satisfying criterion E. If in addition the set $\mathfrak{S}$, i.e., the set of vectors $\Phi[l^a]$ at any one time, is *complete*, then the "classical" action principle (31) is exact.

For exact action functionals, the physical transition matrix element S to go from one state $\Phi[l^a] \in \mathfrak{S}$ at time 0 to another state $\Phi[l'^a] \in \mathfrak{S}$ at time t, has a simple appearance. In particular, from (35), we see that

$$\begin{aligned}S_{l',l} &= (\Phi[l'^a], \Phi[l^a(t)]) \\ &= \mathcal{K}(l'^a; l^a(t)).\end{aligned}$$

Thus for exact action functionals, dynamical transition amplitudes may be read directly out of (23), which, in turn, involves only the projection of the vectors in $\mathfrak{S}$ on the fiducial state Φ_0.

When $\mathcal{H}$ lies outside the Lie algebra, we are generally led to an inexact "classical" action prin-

[15] For a recent discussion of this theorem, see: G. H. Weiss and A. A. Maradudin, J. Math. Phys. 3, 771 (1962).

ciple. Although the equations derived from such an action principle have the form shown in (32), they generally remain inexact for any choice of Φ_0; the exceptional cases, analogous to $\mathfrak{IC}_{h.o.}$ treated in Sec. 2, are discussed below. For reasons similar to those presented in Sec. 2, we may expect to secure classical equations of "maximum accuracy" if we choose Φ_0 so as to extremize $I_{\mathfrak{C}}$. These topics will be discussed elsewhere.

Consideration of Simplifiable, Enlargable, and Special Cases

Suppose now that the Lie algebra is simplifiable in the sense that there exists a choice of infinitesimal elements such that two or more subsets of elements are totally unconnected with one another by the structure constants c^d_{ab}. Then, according to (25) and (26), both M and U provide admixtures only within each individual subset. The canonical kinematical form becomes a sum of terms, each similar to that appearing in (28). It is clear that the Hilbert space in such a case may conveniently be chosen as a product space, a product over as many spaces as there are disconnected subsets of the Lie algebra. If the Hamiltonian is a linear sum of infinitesimal elements, then the Hamiltonian part also breaks up into a sum of terms like (30b), each depending on the parameters within a subset. Such a Hamiltonian, therefore, does not mix the dynamics in one product space with the dynamics in another. The complete problem is a sum of noninteracting smaller problems, one for each of the disconnected subsets of the original Lie algebra.

However, if $\mathfrak{IC}$ is *not* simply a linear sum of infinitesimal elements, then the Hamiltonian defined by Eq. (30a) will possess interaction terms, terms in which the labels from two or more subsets may be involved. In principle, the Lie algebra could be enlarged so as to include $\mathfrak{IC}$ (and generally other elements as well). In this enlarged algebra, with additional parameters, $\mathfrak{IC}$ is now an infinitesimal element and the form in (30b) prevails. Disconnected subsets may be sought in the enlarged algebra. If they are found the problem can be reduced to a sum of simpler noninteracting problems. More specific statements can be made if the Lie algebra were semisimple, for then the disconnected subsets would be a direct sum over simple algebras whose properties are well known.

A particularly simple dynamics arises for those labels belonging to elements in the center $\mathfrak{C}$ of a Lie algebra, i.e., those infinitesimal elements that commute with all other infinitesimal elements. If L_b is such an element, then (25) and (26) state, respectively, that $M^c_b = U^c_b = \delta^c_b$. From (29a) we see that the contribution to (28) of such elements is simply a total differential, $v_b\, dl^b$, summed only over the elements in $\mathfrak{C}$. The contribution of these terms to the Hamiltonian part is trivial in the case (30b), i.e., $H = h^b v_b +$ (terms involving noncentral labels). Thus, the complete dynamics is not fully determined, the evolution of the parameters of the elements in $\mathfrak{C}$ being arbitrary; e.g., if we were to set $\dot{l}^b = h^b$, as suggested by (35), then the appearance of these variables as well as their energy shifts would disappear completely. The basic dynamical elements in this case lie outside the center $\mathfrak{C}$.

It would be possible to use this information regarding the time behavior of the parameters of the elements in $\mathfrak{C}$ to simplify the "classical" action functional. Thus, of the possible paths $\Phi[l^a(t)]$, we might consider only those for which the central element parameters equal specific time functions which satisfy their elementary equations of motion $\dot{l}^b = h^b$; in the derivation of the "classical" equations only the remaining variables would be varied. Care should be taken, however, lest the restricted set of states with only noncentral element parameters free to vary fails to span $\mathfrak{H}$.

If, on the other hand, the Hamiltonian is *not* simply a linear sum of infinitesimal elements, then $H(l^a)$ defined by (30a) may very well depend on the parameters of those elements in the center. Equations generated by extremal conditions for central element parameters are then constraints, $0 = \partial H/\partial l^b$, i.e., the parameters relating to the elements in $\mathfrak{C}$ enter the Lagrangian at most only in the form of Lagrange multipliers.

After establishing that the Hamiltonian $\mathfrak{IC}$ is a member of the Lie algebra, suppose we further find that $\mathfrak{IC}$ lies in a *sub*algebra. Then it may be desirable to simplify the action functional by simply setting $l^a(t) = 0$ (or more generally their values for the identity element) for those infinitesimal elements *outside* the subalgebra. Thus, we are restricting our algebra in such a way that only the subalgebra containing $\mathfrak{IC}$ appears. Such a restriction should be carried out only if (or should be carried out only to an extent that) the vectors remaining in $\mathfrak{S}$ form a complete set. Otherwise the exact "classical" action principle possible in such a case would be restricted to apply to an incomplete set, and the dynamics for an arbitrary state vector could not be predicted.

A special case arises if the Hamiltonian $\mathfrak{IC}$ is an element of a Lie algebra, the remaining elements of which form an invariant subalgebra, and if $\mathfrak{IC}$

has at least one discrete eigenvalue. In such a case, the solution (35) is applicable and we express it as

$$\Phi(t) = \exp(tL_h) \exp[l^a(0)L_a]\Phi_0;$$

the summation over "a" also includes the element L_h, the infinitesimal element representing $\mathfrak{K}$. This solution may be written in the form

$$\Phi(t) = \exp[\bar{l}^a(t)L_a] \exp(tL_h)\Phi_0, \tag{36}$$

where

$$\bar{l}^a(t)L_a \equiv l^a(0)e^{tL_h}L_a e^{-tL_h}.$$

From the definition of $\bar{l}^a$, it is clear that

$$\bar{l}^h(t) = l^h(0), \tag{37a}$$

$$\bar{l}^b(t) = \bar{l}^b[l^c(0), t]; \qquad b, c \neq h. \tag{37b}$$

Now choose Φ_0 to be one of the eigenvectors of $\mathfrak{K}$, hence also of L_h, such that $\exp(tL_h)\Phi_0 = \exp(-i\omega t)\Phi_0$. Then, apart from a trivial phase factor (which may be absorbed if one of the elements L_a is, or effectively acts as unity), the solution (36) has the special form

$$\Phi(t) = \exp[\bar{l}^a(t)L_a]\Phi_0.$$

According to (37), if we choose $l^h(0) = 0$, then $\bar{l}^h(t) = 0$, and furthermore, there is no disturbance to the remaining labels $\bar{l}^b(t)$, $b \neq h$. Thus by this choice, all appearance of the Hamiltonian label l^h can be eliminated from $\mathfrak{S}$, and if the remaining vectors in $\mathfrak{S}$ are a complete set, then the action functional remains exact. It is just this situation that occurred for the harmonic-oscillator example discussed in Sec. 2.

There also may be a simplification in the parameterization when certain of the constants v_d in (27) vanish. From Eq. (29a) it is clear that y_b will in general be simpler if some of the v_d vanish. We shall exclude cases where all the constants v_d vanish, since then the canonical kinematical form itself vanishes. In the elementary example of Sec. 2, for instance, we chose Φ_0 in Eq. (6) so that two out of three such terms would vanish. The vanishing of these expectation values was extended even further to include the Hamiltonian in the case of the harmonic oscillator.

An even greater simplification of the parameterization may take place if the stronger conditions $L_{m_1}\Phi_0 = L_{m_2}\Phi_0 = \cdots = 0$ hold true for a set of elements $\{L_{m_1}, L_{m_2}, \cdots\}$ which form a subalgebra. Let us order our labeling so that L_m, $m = 1, \cdots, M$ denotes the elements in such a subalgebra, and L_p, $p = M + 1, \cdots, N$ denotes the remainder of the elements in the algebra. Then $\Phi[l^a]$ in (22) has the special property that if $l^p = 0$, $p = M + 1, \cdots, N$, then $\Phi[l^m] = \Phi_0$, *independent* of the values of the remaining parameters. This suggests that $\Phi[l^a]$ really depends not on N variables but only on $N - M$ variables. Call these independent variables r^q, $q = 1, \cdots, N - M$. Then in general we expect that there is a many–one mapping of points l^a to points r^q such that $\Phi[l^a] = \Phi[r^q]$. The set of points l^a mapped onto one point r^q may be found from

$$\exp(l^aL_a)\Phi_0 = \exp(l^aL_a) \exp(s^mL_m)\Phi_0 \equiv \exp[\tilde{l}^a(l^b, s^m)L_a]\Phi_0.$$

For fixed l^b the set

$$\mathfrak{W} \equiv \{\tilde{l}^a \mid \tilde{l}^a = \tilde{l}^a(l^b, s^m), s^m \text{ arbitrary}\}$$

is mapped onto a single point whose coordinate values r^q are determined from $N - M$ continuous, linearly independent set functions: $r^q = f^q(\mathfrak{W})$. It is important that the L_m form a subalgebra in the above analysis.

The possibility that $L_m\Phi_0 = 0$ seems to be not an uncommon circumstance as the following example shows. Consider the Lie group $SU(n)$, the n-dimensional unitary–unimodular group, acting on an n-dimensional Hilbert space. Without loss of generality, we can take a representation in which Φ_0 is represented by one in the first row and zero in the remaining rows. Now there exists a subgroup of $SU(n)$ that leaves Φ_0 invariant, and this subgroup is clearly isomorphic to $SU(n - 1)$. Thus of the $n^2 - 1$ parameters in $SU(n)$, a number corresponding to $SU(n - 1)$, i.e., $(n - 1)^2 - 1$ are totally arbitrary. This leaves $(n^2 - 1) - (n - 1)^2 + 1 = 2n - 1$ effective parameters to describe Φ. While an expression of these effective parameters in terms of those of $SU(n)$ is, in general, very complicated, it is easy to see that the *number* $2n - 1$ is correct, since, in a complex n-dimensional space, there are $2n$ real variables needed to describe a vector, less one to account for normalization.

It is worth speculating at this point that an extension of the analysis of those cases where several L_m annihilate Φ_0 may shed some light on the form taken in quantum mechanics by classical guage groups. Although our present analysis is basically relevant to a finite number of degrees of freedom, it certainly contains non-Abelian classical c-number symmetries. Thus, the introduction of infinitely many similar spaces to describe a field, and the enlargement of the symmetries to describe locally variable guages may well clarify the quantum treatment of such questions. We hope to comment on this possibility in subsequent work.

An Example

It is clear that one example of the formalism developed in this section is the one-dimensional problem treated in Sec. 2. We quote without proof some results of another application of our general formalism to a two-dimensional Hilbert space. We choose as infinitesimal elements of our Lie algebra $\frac{1}{2}i\boldsymbol{\sigma}$, where $\boldsymbol{\sigma}$ are the three Pauli spin matrices. A vector notation will be used throughout to treat the label indices. The members of the OFS are $\Phi[\mathbf{l}] \equiv \exp(\frac{1}{2}i\mathbf{l}\cdot\boldsymbol{\sigma})\Phi_0$. The fiducial state is characterized as in (27) by giving $\mathbf{v} \equiv (\frac{1}{2}\hbar)(\Phi_0, \boldsymbol{\sigma}\Phi_0)$. In terms of these quantities, the canonical kinematical form reads

$$i\hbar(\Phi, d\Phi) = \mathbf{y}\cdot d\mathbf{l},$$

where

$$\mathbf{y} \equiv (l^{-3}\sin l - l^{-2})(\mathbf{l}\cdot\mathbf{v})\mathbf{l} - l^{-1}\sin l\,\mathbf{v} + l^{-2}(1 - \cos l)\mathbf{l}\times\mathbf{v}.$$

Here l denotes the magnitude of $\mathbf{l}$. If we choose a Hamiltonian of the form $(\frac{1}{2}\hbar)\mathbf{h}\cdot\boldsymbol{\sigma}$ [cf. Eq. (30b)], then an exact "classical" action principle results whose Hamiltonian is

$$H(\mathbf{l}) = \cos l(\mathbf{h}\cdot\mathbf{v}) + l^{-2}(1 - \cos l)(\mathbf{h}\cdot\mathbf{l})(\mathbf{v}\cdot\mathbf{l}) + l^{-1}\sin l(\mathbf{l}\times\mathbf{h}\cdot\mathbf{v}).$$

We shall not pursue the resulting equations of motion, save to remark that the evolution of $\mathbf{l}$ is necessarily nonunique since the Lie algebra has odd dimensionality.

The introduction of new labels other than the "canonical coordinates" we have been using cannot change the physics of a given problem but only its description. We now wish to point out that quite different labels give to the preceding example a much simpler appearance. For this purpose we choose to label an equivalent OFS by Eulerian angles:

$$\Phi[\theta, \varphi, \psi] = \begin{pmatrix} e^{-\frac{1}{2}i(\psi+\varphi)}\cos\frac{1}{2}\theta \\ e^{-\frac{1}{2}i(\psi-\varphi)}\sin\frac{1}{2}\theta \end{pmatrix}.$$

In these variables, the canonical kinematical form is expressed by

$$i\hbar(\Phi, d\Phi) = (\tfrac{1}{2}\hbar)\cos\theta\, d\varphi + (\tfrac{1}{2}\hbar)\, d\psi,$$

and the "classical" Hamiltonian becomes $(\frac{1}{2}\hbar)\mathbf{h}\cdot\mathbf{w}$, where

$$\mathbf{w} \equiv (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta).$$

While the unit vector $\mathbf{w}$ appears to be an ordinary three-vector, the unusual role of $(\frac{1}{2}\hbar)\cos\theta$ as a momentum conjugate to φ can be shown to infer the Poisson bracket relation $[w_x, w_y]_{\text{P.b.}} = -2w_z/\hbar$ characteristic of an angular momentum. Thus, spin degrees of freedom in our formalism have a "classical" description of the *form* but not the *interpretation* extensively discussed by Bohm and coworkers.[16]

Resolutions of Unity and Continuous Representations

For completeness, let us make some remarks regarding resolutions of unity expressed in terms of the vectors $\Phi[l^a]$, which we assume to satisfy Postulate 3 of I, i.e., so that the form

$$1 = \int \Phi[l^a]\delta l^a \Phi^\dagger[l^a] \tag{38}$$

is true. Since the $V[l^a]$ coefficients in (22) form a group, they are also the elements of the invariance group $\mathcal{G}$ of $\mathfrak{S}$. As such they form a transitive group in which any vector can be transformed into any other vector in $\mathfrak{S}$. When $\mathcal{G}$ is compact, the invariant measure theorem, Theorem 1 of I, assures us that if (38) is true, we may without loss of generality choose δl^a as the invariant group measure. Indeed, when the representation of the $V[l^a]$ is irreducible, Schur's lemma guarantees (38) for any Φ_0. If the group $\mathcal{G}$ is not compact, a possible candidate for δl^a is still the (left-) invariant group measure. If the Lie algebra is simplifiable in our earlier sense, then the resolution of unity is a product resolution over each of the product spaces that make up the Hilbert space $\mathfrak{H}$.

The assumed validity of (38) gives rise to a representation of $\mathfrak{H}$ by means of continuous label-space functions. In particular such functions are

$$\psi(l^a) \equiv (\Phi[l^a], \Psi),$$

while from (38) the inner product takes the form

$$(\Psi_1, \Psi_2) = \int \psi_1^*(l^a)\delta l^a \psi_2(l^a). \tag{39}$$

The functions $\psi(l^a)$ representing vectors in Hilbert space are not arbitrary but must satisfy the projection identity

$$\psi(l'^a) = \int \mathcal{K}(l'^a; l^a)\psi(l^a)\delta l^a, \tag{40}$$

where $\mathcal{K}$ is defined in (23). The solutions to (40) form a linear vector space and include as special cases those functions which represent vectors in $\mathfrak{S}$. That is, when $\Psi = \Phi[l''^b]$, Eq. (40) holds for the special case $\psi(l^a) = \mathcal{K}(l^a; l''^a)$ [cf. Eq. (6) of I].

[16] See, e.g., D. Bohm, R. Schiller, and J. Tiomno, Nuovo Cimento Suppl. 1, 48 (1955).

Such a constraint on $\mathcal{K}$ used in conjunction with the canonical kinematical form (28) could be helpful in *deriving* $\mathcal{K}$ starting purely from the "classical" theory.

4. LABELING BY VECTOR COMPONENTS

Let us confine our attention initially to an n-dimensional Hilbert space, but allow any Hermitian operator to be a potential candidate for the Hamiltonian. To ensure that we obtain an exact "classical" action principle, we shall parameterize and include all unit vectors in our overcomplete family of states. For example, the parameters of the Lie group $SU(n)$ could be used and even arranged so that just one such parameter was associated with the Hamiltonian. However, many of the remaining parameters would be superfluous, and the vectors would be independent of them. A symmetric and virtually nonredundant set of parameters may be introduced in the following manner.

Let $\psi = \psi_1, \cdots, \psi_n$ be an n-tuple of complex numbers lying on the complex n-dimensional unit sphere S_c^n;

$$|\psi|^2 \equiv \sum_{k=1}^{n} |\psi_k|^2 = 1. \tag{41}$$

Then each point in S_c^n corresponds to a unit vector in Hilbert space and, by means of a suitable mapping $\mathfrak{M}$: $\psi \rightarrow \Phi[\psi]$, we can characterize each unit vector by a "label" ψ. The inner product of two such vectors can be expressed as a function of their labels, which we shall define as

$$(\Phi[\psi'], \Phi[\psi]) \equiv \psi'^*\psi \equiv \sum_k \psi_k'^*\psi_k. \tag{42}$$

Postulate 1 of I, regarding the local density of the vectors in $\mathfrak{S}$ as well as the completeness aspect of Postulate 3 of I are trivially fulfilled. The continuity of the labeling, Postulate 2 of I, is satisfied in virtue of the form adopted in (42).

If we now consider vector-valued time functions we put as before $\Phi(t) = \Phi[\psi(t)]$. The canonical kinematical form follows from the differential of (42) as

$$i\hbar(\Phi, d\Phi) = i\hbar\psi^* \, d\psi. \tag{43}$$

The "classical" Hamiltonian must be a bilinear functional in ψ^* and ψ and is of the general form

$$H = (\Phi[\psi], \mathfrak{K}\Phi[\psi]) \equiv \psi^*\tilde{\mathfrak{K}}\psi. \tag{44}$$

The equations of motion that follow from an action principle based on (43) and (44) are, as expected, $i\hbar \, \partial\psi/\partial t = \tilde{\mathfrak{K}}\psi$. Since the solution must remain a unit vector in the Hilbert space, it is characterized by some label ψ, and consequently the "classical" action functional is exact. This example is one where the generalized "classical" theory, as we have defined the term, will contain the same physics and very nearly the same formalism as the quantum mechanics. Here the relative term "classical" is interchangeable with the term "quantum."

As regards the resolution of unity, we can expect the following form to hold:

$$1 = \int \Phi[\psi] \, d\bar{\mu}\,(\psi)\Phi^\dagger[\psi]. \tag{45}$$

The set $\mathfrak{S}$ is invariant under any and all unitary transformations. If we invoke the invariant measure theorem, Theorem 1 of I, then Schur's lemma insures that the integral in (45) is in fact necessarily proportional to the unit matrix. The appropriate measure on vectors has the symmetric form

$$d\bar{\mu}(\psi) = \left(\frac{n!}{\pi^n}\right)\delta(1 - \sum_k |\psi_k|^2) \prod_k d\psi_{kr} \, d\psi_{ki}, \tag{46}$$

where r and i denote the real and imaginary parts, respectively.

From (45) there arises a representation of Hilbert-space vectors by functions of ψ^* homogeneous in the first degree. The inner product of two vectors is then expressed by

$$(\Psi_1, \Psi_2) = \int \omega_1^*(\psi) \, d\bar{\mu}\,(\psi)\omega_2(\psi),$$

where

$$\omega(\psi) \equiv (\Phi[\psi], \Psi) \equiv \sum_{k=1}^{n} \psi_k^*\nu_k,$$

and ν_k are n complex coefficients.

Two-Dimensional Space

Consider the case $n = 2$. Of the four real parameters in ψ_1 and ψ_2, one may be eliminated by means of the constraint (41) and another represents an over-all phase factor which can not be a true dynamical variable. [The over-all phase was useful in establishing the weight factor (46) in the resolution of unity, but may now be eliminated.] Thus there are only two dynamical degrees of freedom. In order to more clearly display these two degrees of freedom, we proceed as follows.

Let N be a projection operator with eigenvectors $\Phi^{(0)}$, and $\Phi^{(1)}$ such that $N\Phi^{(r)} = r\Phi^{(r)}$. We now define

$$N\Phi[\psi] = (1 - |\chi|^2)^{\frac{1}{2}}e^{-i\alpha/\hbar}\Phi^{(1)}, \tag{47a}$$

where χ is a complex variable restricted so that $0 \leq |\chi| \leq 1$, and $-\alpha/\hbar$ is the phase of the projected component. To fix the phase of χ, we set

$$(1 - N)\Phi[\psi] = \chi e^{-i\alpha/\hbar}\Phi^{(0)}. \tag{47b}$$

When expressed in these variables, the canonical kinematical form becomes

$$i\hbar(\Phi, d\Phi) = (i\hbar/2)(\chi^* \overleftrightarrow{\partial}_t \chi)\, dt + d\alpha, \tag{48}$$

where $A \overleftrightarrow{\partial}_t B \equiv A(\partial B/\partial t) - (\partial A/\partial t)B$. The Hamiltonian is clearly a function only of χ and χ^* defined by

$$H(\chi) = (\Phi[\psi], \mathcal{H}\Phi[\psi]). \tag{49}$$

The total differential $d\alpha$ will not effect the dynamics, and may be arbitrarily specified. As an example, therefore, let us restrict $\mathfrak{S}$ so as to include only those vectors of the form (47) *for which* $\alpha = 0$. Now that we have restricted our OFS to only those vectors parameterized by true dynamical degrees of freedom, the question arises whether there remain any Hamiltonians for which the "classical" action principle will be exact.

To answer this question we observe that any vector that we choose as an initial vector will have a real coefficient of $\Phi^{(1)}$. As time progresses, this property must be maintained for any choice of χ at $t = 0$; hence there can be no mixing of (47a) and (47b) as time passes. The state $\Phi^{(1)}$ must be an eigenvector of the Hamiltonian with eigenvalue zero. Thus the most general solution is $\mathcal{H} = \hbar\omega(1 - N)$, which in turn leads to

$$H(\chi) = \hbar\omega\chi^*\chi. \tag{50}$$

We see that the "classical" action principle

$$I = \int (\tfrac{1}{2}i\hbar\chi^* \overleftrightarrow{\partial}_t \chi - \hbar\omega\chi^*\chi)\, dt \tag{51}$$

for a single fermion oscillator is an exact action principle. This conclusion remains true even if ω is an explicit function of t. Furthermore, the interpretation of χ follows from (50); as usual, it is simply a probability amplitude for oscillator excitation.

Generalization to a Fermion Field

An infinite linear sum of action functionals of the type in (51), each characterizing an independent fermion oscillator, can be used to describe a fermion field. So long as the oscillators remain independent, the over-all action functional will be exact. It follows, for example, that the conventional Dirac c-number action functional in the presence of an external, fixed c-number source may be considered as an exact "classical" action principle since it may always be resolved into independent noninteracting oscillators. If the source is also allowed to respond dynamically it means, in general, that the "classical" action is no longer exact. In either case the "classical" action functional arises as a restricted evaluation of that action, of the general form in Eq. (1), which leads to the so-called second quantized Schrödinger equations.[17] Thus we find the satisfying result that the Dirac equations are simply "classical" equations relative to the second quantized formalism for fermion fields, exact in the absence of dynamical interactions, with all the Fermi–Dirac statistics being correctly included by the limitation $0 \leq |\chi| \leq 1$ placed on the "classical" amplitude of each independent oscillator.

5. SUMMARY

We have focused our attention in this paper on the relation of quantum and classical dynamics from the standpoint of continuous-representation theory and its associated overcomplete families of states. A study of the elementary examples in Sec. 2 suggested that classical mechanics can already be viewed as the study of quantum mechanics for a certain restricted class of vectors if only we reinterpret the classical variables as labels for those vectors. The generalization of this result led to our concept of "classical" dynamics relative to a set $\mathfrak{S}$ as the study of quantum dynamics for unit vectors restricted to the subset $\mathfrak{S}$. Such a definition for the restricted dynamics merits the name "classical" since it deals with c-number variables capable of continuous variation with the aid of conventional action principle techniques.

In a larger and more abstract sense it should be recognized that the dynamical variables are invariantly characterized as the Hilbert space vectors themselves, it being expedient to discuss these vectors by the labels we introduce.

There are several aspects of our formalism and viewpoint worth noting. Firstly, the construction and analysis of classical theories becomes at the same time a partial study of quantum-mechanical theories. For simple enough systems we have learned that the classical dynamics is sufficient to infer the correct quantum dynamics.

[17] For further details relating to the evaluation of the action principle for the relevant restricted set of states, see J. R. Klauder, Ann. Phys. (NY) **11**, 123 (1960), pp. 159 and 160. Many of the formal manipulations in that reference regarding the measure on label-space points may be eliminated by the conventional device of first working in a "box" of finite volume and later passing to the limit.

Secondly, we have seen that the appropriate "classical" theory is dictated once we are given the subset $\mathfrak{S}$ of Hilbert space vectors, the labels for these vectors, and the Hamiltonian. Thus the possible forms of "classical" action functionals can be classified and catalogued, e.g., in the manner discussed in Sec. 3. Then, were we confronted with a specific classical theory, the possible associated quantum theories and their properties could be readily determined, at least in principle.

Thirdly, let us reconsider the "process of quantization" from our viewpoint. Initially suppose we are given a classical theory in the form of an action functional expressed in terms of c-number dynamical variables. The first step is to reinterpret the classical variables as vector labels and to view the action functional as a restricted evaluation of the true quantum action functional. It is in this step that the conventional factor-ordering ambiguity, if any, would show up. For example, given only $\mathfrak{IC}(p, q)$ in Eq. (10) the term $\mathcal{O}$ must first be chosen before the "classical" Hamiltonian $H(p, q)$ is determined. While some of the freedom in $\mathcal{O}$ stems from the arbitrariness in Φ_0, some may also lie in the factor ordering in $\mathfrak{IC}$. It is our contention that the proper choice of the classical Hamiltonian should already coincide with one of the possible expectation values H; the general separation of the classical Hamiltonian into $\mathcal{O}$ and $H - \mathcal{O}$ is to be regarded as heuristic and not of fundamental significance. Adopting this point of view, our first step then neither changes the form of the classical action functional nor alters the mathematical properties of the dynamical variables (e.g., c-number $\rightarrow$ operator); it is strictly a *reinterpretation* of the physical meaning of the old c-number variables.

The second step of quantization involves an *enlargement*, in one way or another, of the *domain* of the action functional so as to infer the true quantum dynamics. As we have seen, this enlargement need only proceed to a point where an exact "classical" action functional arises. Independent of just how far this domain enlargement proceeds, the process of quantization, which—as far as the dynamical formalism is concerned—is entirely contained in this domain expansion, is seen to involve a smooth and continuous transition.[18] This desirable conceptual feature, coupled with the universal applicability of our approach to any dynamical system, and coupled with the physically desirable correspondence that eliminates the zero-point field energies, all provide strong reasons to favor our view of "classical" theories as simply being restricted quantum theories.

---

[18] An alternate means to pass from a classical to a quantum theory is by means of the Feynman sum-over-histories. The analogue of this technique in our formalism has a somewhat different form than the usual one; it is discussed formally in general terms similar to those of the present paper in J. R. Klauder, Ann. Phys. (NY) **11**, 123 (1960), pp. 142–149, and in unpublished lecture notes "The Sum-Over-Histories: Formalism and Some Applications," University of Bern, Switzerland, 1962.

A related formulation of the Schwinger Action Principle approach to quantum mechanics that is suitable only for infinite-dimensional Hilbert spaces is discussed in reference 3. A more general statement of the Schwinger Action Principle is implicitly contained in our formalism. For example, the basic kinematical effects are contained in an expression of $\delta(\Phi[l^{a'}], \Phi[l^a])$ in terms of label differentials with the aid of the formulas in Sec. 3.

A Thermodynamical Limitation on Compressibility*

LINCOLN E. BRAGG AND BERNARD D. COLEMAN
Mellon Institute, Pittsburgh, Pennsylvania
(Received 20 February 1963)

In their theory of thermostatics, Coleman and Noll have obtained a convexity inequality which places restrictions on admissible stress–strain functions for elastic materials. Here we show that for an arbitrary elastic material in an arbitrary state of strain F, the general convexity inequality implies that the modulus of compression k obeys the inequality $k(F) \geq \frac{2}{3}\bar{p}(F)$ where $\bar{p}$ is the mean pressure, i.e. minus one-third the sum of the principal stresses. Here k is defined to be the derivative of $\bar{p}$ with respect to the mass density along a deformation process representing a uniform expansion from the state F.

INTRODUCTION

IN the classical theory of infinitesimal deformations about a stress-free state of an isotropic elastic solid, it is usually assumed that the Lamé coefficients λ and μ obey inequalities

$$\lambda + \tfrac{2}{3}\mu > 0, \tag{1a}$$

$$\mu > 0, \tag{1b}$$

which state that the modulus of compression $k = \lambda + \frac{2}{3}\mu$, and the shear modulus μ both be positive. These inequalities are not only intuitively obvious but also mathematically important, for upon them rest uniqueness theorems for boundary-value problems in the classical theory. Yet the ease with which the physical intuition accepts the inequalities (1) becomes perplexing when one goes into the matter deeply. Indeed, the problem of finding the generalization of (1a) and (1b) applicable to *finite deformations of arbitrary elastic materials* appears to be unsolved. The present article is concerned with a generalization of (1a).

In their theory of the thermostatics of continuous media,[1,2] Coleman and Noll have laid down postulates which lead to a general convexity inequality restricting the form of stress–strain functions for elastic materials. Here we call that inequality the TI (Thermostatical Inequality). Consequences of the TI are known for various types of materials. In the theory of *infinitesimal* deformations from a natural state of an isotropic solid, the TI reduces to the inequalities (1). All the known consequences of the TI for *finite* deformations seem to be in accord with physical experience in solids.

For elastic fluids (i.e. materials for which the stress is always a hydrostatic pressure depending only on the density) the TI is equivalent to the assertions that the pressure p be positive for each value of the density ρ and the modulus of compression $\rho\, dp(\rho)/d\rho$ be greater than two-thirds of the pressure[3]:

$$p(\rho) > 0, \tag{2a}$$

$$\rho\, dp(\rho)/d\rho \geq \tfrac{2}{3}p(\rho). \tag{2b}$$

In order for (2a) and (2b) to be sufficient, as well as necessary, for the TI in fluids it must be understood that equality occurs in (2b) on at most a nowhere-dense set of values of ρ; such subtleties do not interest us here, however.

That the pressure is positive seems to be in accord with experience. Yet, since there can be, near the critical point, a range of densities at which the modulus of compression is less than two-thirds of the pressure, the TI does not apply to all fluids in all circumstances.

The TI is equivalent to a requirement of stability against homogeneous disturbances at fixed surface forces. Heuristically, such a requirement appears appropriate for solids, because it is surface forces which are controlled in most mechanical measurements on solids. For fluids it is usually the pressure that is controlled, and the surface forces, instead of remaining fixed, must then change their direction to stay normal to the surface and their magnitude to compensate for alteration in surface area. Thus we cannot expect the TI to be applicable to fluids, and it is still an open question whether it is possible to find a single inequality which yields all thermodynamical restrictions on the local static behavior of elastic materials, whether they be fluids, solids, or neither.

Since we believe the TI to be physically applicable

* The research reported here was supported by the Air Force Office of Scientific Research under Contract AF 49(638)541.

[1] B. D. Coleman and W. Noll, Arch. Rational Mech. Anal. 4, 97 (1959).

[2] For alternative presentations and extensions of the theory, see reference 3.

[3] B. D. Coleman, Arch. Rational Mech. Anal. 9, 172 (1962).

to solids rather than fluids, we here attempt to derive from it a generalization of (2b) which is broad enough to apply to solids; in doing this we obtain a generalization of (1a) meaningful for arbitrary solids, whether isotropic or not, in arbitrary states of strain.

1. KINEMATICAL PRELIMINARIES

Let X be a material point of a body $\mathcal{B}$. Let $\mathbf{x}$ be the position in space of X corresponding to some configuration $\mathcal{R}$ of $\mathcal{B}$ which we take to be a reference configuration. Let $\mathbf{y}$ be the position in space of X for any other configuration $\mathcal{C}$. Keeping $\mathcal{C}$ and $\mathcal{R}$ fixed, as we vary X over $\mathcal{B}$, we obtain different values of $\mathbf{x} = \mathbf{x}(X)$ and $\mathbf{y} = \mathbf{y}(X)$. Since these values are in one-to-one correspondence we can regard $\mathbf{y}$ as a function of $\mathbf{x}$;

$$\mathbf{y} = \mathbf{y}(\mathbf{x}). \tag{1.1}$$

The gradient F of this function,

$$F = \nabla \mathbf{y}(\mathbf{x})|_{\mathbf{x}=\mathbf{x}(X)}, \tag{1.2}$$

is called the *deformation gradient* (at X) of the configuration $\mathcal{C}$ relative to the reference configuration $\mathcal{R}$. We note that this second-order tensor[4] F depends on both $\mathcal{C}$, $\mathcal{R}$ and, in general, the material point X. Since we assume that $\mathbf{y}(\mathbf{x})$ is always a one-to-one smooth function with a smooth inverse, F is always an invertible tensor, i.e. has an inverse F^{-1} such that $F^{-1}F = FF^{-1} = I$, where I is the unit tensor. Hence, the determinant of F, det F, is nonzero.

If $\rho_{\mathcal{C}}$ and $\rho_{\mathcal{R}}$ are the mass densities at X corresponding to the configurations $\mathcal{C}$ and $\mathcal{R}$, then

$$\rho_{\mathcal{C}}^{-1} = |\det F|\, \rho_{\mathcal{R}}^{-1}. \tag{1.3}$$

If F is a proper orthogonal tensor, then we say that X is *rigidly rotated* when the configuration of $\mathcal{B}$ is changed from $\mathcal{R}$ to $\mathcal{C}$. If, on the other hand, F is positive-definite and symmetric, then we say that X experiences a *pure stretch* on going from $\mathcal{R}$ to $\mathcal{C}$.

If F_0 and F_1 are the deformation gradients, at X, of $\mathcal{C}_0$ and $\mathcal{C}_1$, respectively, relative to the same fixed reference configuration $\mathcal{R}$, then the deformation gradient, at X, of $\mathcal{C}_1$ relative to $\mathcal{C}_0$ is the tensor $F_1F_0^{-1}$.

2. THERMOSTATICAL INEQUALITIES

We say that the material at X is an *elastic material* if the stress S at X (in the configuration $\mathcal{C}$) is determined by the deformation gradient F at X:

$$S = \mathcal{S}(F). \tag{2.1}$$

Of course, the form of the function $\mathcal{S}$ in (2.1) depends on the choice of the reference configuration $\mathcal{R}$.

In general theories of thermostatics it is recognized that the stress S can depend not only on the deformation gradient F, i.e. on the "strain," but also on a thermodynamical parameter such as the temperature θ or the entropy density η. Our present formulas and inequalities are to be regarded, physically, as pertaining to situations in which either θ or η is controlled and held at a fixed value.

The axioms for thermostatics laid down in reference 1 are shown in reference 3 to yield the following assertion:

Thermostatical Inequality (TI): Consider the class $\mathfrak{C}$ of smooth curves F_t with values in the space of all invertible tensors F, and let the parameter t for these curves vary from 0 to 1. Let $\mathfrak{C}'$ be the set of all curves in $\mathfrak{C}$ for which $F_0 \neq F_1$ and $F_1F_0^{-1}$ is both symmetric and positive-definite. Then, the following inequality must hold for all curves in $\mathfrak{C}'$,[5]

$$\rho_0^{-1} \int_0^1 |\det F_t|\, \operatorname{tr}\left\{F_t^{-1}\mathcal{S}(F_t)\,\frac{dF_t}{dt}\right\} dt > \rho_0^{-1}\,|\det F_0|\, \operatorname{tr}\{F_0^{-1}\mathcal{S}(F_0)[F_1 - F_0]\}. \tag{2.2}$$

Here tr $\{\ \}$ is the trace operation. The integral on the left is to be interpreted as a line integral along F_t from $t = 0$ to $t = 1$.

Remark: If we let ρ_0 represent the density corresponding to a configuration with deformation gradient F_0, then the quantities appearing on each side of (2.2) represent work, per unit mass, done against contact forces at a material point X as the local configuration about X is deformed from one with deformation gradient F_0 to one with deformation gradient F_1 along the path described by F_t. The quantity on the left in (2.2) gives the "true" work done, i.e. the work done assuming that at each t the contact forces on each material surface at X are those which one calculates using the stress tensor $S = \mathcal{S}(F_t)$ and the actual configuration of the surface at t; the quantity on the right in (2.2), however, gives the work which would be done along the path F_t if the contact forces were to remain fixed at their initial values.[6] Thus, (2.2) states that $\mathcal{S}$ must be such that contact forces always change in a process which results in a pure stretch, and, further-

[4] We denote tensors by light face Latin majuscules, reserving the symbol X, however, for material points. If the reader wishes, he may regard symbols such as F as representing 3×3 matrices of Cartesian components.

[5] In writing (2.2) we have made use of the fact that $S = \mathcal{S}(F)$ is a symmetric tensor, this enables us to eliminate the many transpositions which occur in Eqs. (1.4) and the inequality (2.3) of reference 3.

[6] Again we note that, when the configuration of a surface is changing, keeping contact forces fixed is not equivalent to keeping the stress tensor fixed.

more, they always change in such a way that the work done against them is greater than that which would have been done had they remained fixed.

Let us consider the class $\mathfrak{C}''$ of curves in $\mathfrak{C}$ which have the form

$$F_t^{(1)} = A + Bt, \qquad 0 \leq t \leq 1,$$

where A and B are invertible tensors, and are such that

$$F_1^{(1)}F_0^{(1)-1} = (A + B)A^{-1}$$

is positive-definite and symmetric. Clearly all the curves in $\mathfrak{C}''$ are also in $\mathfrak{C}'$. Furthermore, for each curve $F_t^{(1)}$ in $\mathfrak{C}''$, there is another curve $F_k^{(2)}$ which can be written in the form

$$F_t^{(2)} = A + B - Bt, \qquad 0 \leq t \leq 1,$$

and which represents $F_t^{(1)}$ "traversed in the opposite sense." For $F_t^{(2)}$, we have

$$F_1^{(2)}F_0^{(2)} = A(A + B)^{-1} = [(A + B)A^{-1}]^{-1},$$

which is positive-definite and symmetric if and only if $(A + B)A^{-1}$ is. Hence $F_t^{(2)}$ is in $\mathfrak{C}''$ and thus in $\mathfrak{C}'$. Therefore, it follows from the TI that (2.2) must continue to hold if F_t in (2.2) is replaced by either $F_t^{(1)}$ or $F_t^{(2)}$. On adding the two inequalities so obtained and noting that

$$\rho_0^{-1}\int_0^1 |\det F_t^{(1)}|\,\operatorname{tr}\left\{F_t^{(1)-1}\mathcal{S}(F_t^{(1)})\frac{dF_t^{(1)}}{dt}\right\}dt$$
$$= -\rho_0^{-1}\int_0^1 |\det F_t^{(2)}|\,\operatorname{tr}\left\{F_t^{(2)-1}\mathcal{S}(F_t^{(2)})\frac{dF_t^{(2)}}{dt}\right\}dt,$$

we obtain the inequality

$$tr\{(F_{(2)} - F_{(1)})[\rho_{(2)}^{-1}F_{(2)}^{-1}\mathcal{S}(F_{(2)}) - \rho_{(1)}^{-1}F_{(1)}^{-1}\mathcal{S}(F_{(1)})]\} > 0; \qquad (2.3)$$

here we have put

$$F_{(1)} = F_0^{(1)} = A, \qquad F_{(2)} = F_0^{(2)} = A + B,$$
$$\rho_{(1)}^{-1} = |\det F_{(1)}|\,\rho_0^{-1}, \quad \rho_{(2)}^{-1} = |\det F_{(2)}|\,\rho_0^{-1};$$

hence,

$$\rho_2^{-1} = |\det (F^{(2)}F^{(1)-1})|\,\rho_1^{-1}. \qquad (2.4)$$

On noticing that $F_{(2)}F_{(1)}^{-1} = (A + B)A^{-1}$, we see that the inequality (2.3) must hold for all invertible tensors $F_{(1)}, F_{(2)}$ such that $F_{(1)} \neq F_{(2)}$ and $F_{(2)}F_{(1)}^{-1}$ is positive-definite and symmetric.

We call the inequality (2.3) the Weakened Thermostatical Inequality (WTI).

We have just proved that the WTI follows from the TI; i.e., every stress–strain function $\mathcal{S}$ compatible with the theory of references 1 and 3 must obey (2.3) for all $F_{(1)}$ and $F_{(2)}$ such that $F_{(2)}F_{(1)}^{-1}$ is positive-definite, symmetric, and $\neq I$.

Remark: The converse of the above result is not true: the WTI does not imply the TI. This is not surprising; for the TI asserts that (2.2) must hold for *all* curves in $\mathfrak{C}'$, while in proving the WTI we used only the assertion that (2.2) hold for *straight lines* in $\mathfrak{C}'$.

In the rest of this article we assume the WTI as an axiom. Hence all the propositions we prove are consequences of the TI.

3. MEAN PRESSURE

If S has the form

$$S = -pI, \qquad (3.1)$$

then we say that the stress is a hydrostatic pressure. If, for some F, $S = \mathcal{S}(F)$ does not have the form (3.1), S can still be uniquely decomposed into a pressure part $-\bar{p}I$ and a deviator T as follows:

$$S = \mathcal{S}(F) = -\bar{p}(F)I + T(F); \qquad (3.2)$$

here $\bar{p}(F)$ is defined by

$$\bar{p}(F) = -\tfrac{1}{3}\operatorname{tr}\mathcal{S}(F), \qquad (3.3)$$

and is called the *mean pressure*. It follows from Eqs. (3.2) and (3.3) that

$$\operatorname{tr} T(F) = 0. \qquad (3.4)$$

We now prove

Proposition 1: For all elastic materials, the mean pressure $\bar{p}(F)$, defined in (3.3), has the property that for each F, $\nu^2\bar{p}(\nu F)$ is strictly decreasing in ν for all $\nu > 0$.

Proof: We use the WTI, putting $F_{(1)} = \nu_{(1)}F$ and $F_{(2)} = \nu_{(2)}F$, with $\nu_{(1)} \neq \nu_{(2)}$, $\nu_{(1)} > 0$, $\nu_{(2)} > 0$. Since $F_{(1)} \neq F_{(2)}$ and $F_{(2)}F_{(1)}^{-1} = (\nu_{(2)}/\nu_{(1)})I$ is positive-definite and symmetric, (2.3) holds and may be written in the form

$$\operatorname{tr}\{(\nu_{(2)} - \nu_{(1)})[\rho_{(2)}^{-1}\nu_{(2)}^{-1}\mathcal{S}(\nu_2 F) - \rho_{(1)}^{-1}\nu_{(1)}^{-1}\mathcal{S}(\nu_{(1)}F)]\} > 0. \qquad (3.5)$$

By definition,

$$\operatorname{tr}\mathcal{S}(\nu_{(1)}F) = -3\bar{p}(\nu_{(1)}F), \quad \operatorname{tr}\mathcal{S}(\nu_{(2)}F) = -3\bar{p}(\nu_{(2)}F), \qquad (3.6)$$

and by (2.4) we have

$$\rho_{(2)}^{-1} = |\det (I\nu_{(2)}/\nu_{(1)})|\,\rho_{(1)}^{-1},$$

i.e.

$$\rho_{(1)}/\rho_{(2)} = (\nu_{(2)}/\nu_{(1)})^3. \qquad (3.7)$$

Hence, (3.6) yields the inequality

$$-(\nu_{(2)} - \nu_{(1)})[\nu_{(2)}^2\bar{p}(\nu_2 F) - \nu_{(1)}^2\bar{p}(\nu_{(1)}F)] > 0.$$

In other words, whenever we have $\nu_{(2)} > \nu_{(1)}$, we must have

$$\nu_{(2)}^2\bar{p}(\nu_{(2)}F) < \nu_{(1)}^2\bar{p}(\nu_{(1)}F), \qquad \text{Q.E.D.}$$

4. MODULUS OF COMPRESSION

Let us now consider the quantity $k(F)$ defined by

$$k(F) = -(\partial/\partial\alpha)\bar{p}(\alpha^{\frac{1}{3}}F)|_{\alpha=1}. \tag{4.1}$$

It follows from (1.3) that if F is the deformation gradient of $\mathcal{C}$ relative to $\mathcal{R}$ then F can be written in the form

$$F = H(\rho_{\mathcal{R}}/\rho_{\mathcal{C}})^{\frac{1}{3}}, \tag{4.2a}$$

where H is unimodular, i.e.,

$$\det H = 1. \tag{4.2b}$$

Equations (4.1) and (4.2) yield

$$k(F) = \rho_{\mathcal{C}}\,\partial\bar{p}(H(\rho_{\mathcal{R}}/\rho)^{\frac{1}{3}})/\partial\rho|_{\rho=\rho_{\mathcal{C}}}. \tag{4.3}$$

In other words, $k(F)$ is a *modulus of compression*. It gives the ratio $\rho\ \partial\bar{p}/\partial\rho$ of the increment $d\bar{p}$ in mean pressure to the relative increment $d\rho/\rho$ in density required to effect a small uniform expansion from the state of strain characterized by F.[7]

Proposition 2: For all elastic materials and for all deformation gradients F,

$$k(F) \geq \tfrac{2}{3}\bar{p}(F). \tag{4.4}$$

Proof: By Proposition 1, for each F, $\alpha^{\frac{1}{3}}\bar{p}(\alpha^{\frac{1}{3}}F)$ is a strictly decreasing function of $\alpha^{\frac{1}{3}}$ for $\alpha^{\frac{1}{3}} > 0$, hence it is a strictly decreasing function of α for $\alpha > 0$, and

$$0 \geq (\partial/\partial\alpha)[\alpha^{\frac{1}{3}}\bar{p}(\alpha^{\frac{1}{3}}F)]|_{\alpha=1}$$

$$= [\tfrac{2}{3}\alpha^{-\frac{1}{3}}\bar{p}(\alpha^{\frac{1}{3}}F) + \alpha^{\frac{1}{3}}(\partial/\partial\alpha)\bar{p}(\alpha^{\frac{1}{3}}F)]_{\alpha=1}. \tag{4.5}$$

On substituting the definition (4.1) into (4.5) we obtain (4.4) immediately, Q.E.D.

Special Cases

Undistorted states of general isotropic materials: If F describes an undistorted configuration of an isotropic material, then $\mathcal{S}(F)$ is a hydrostatic pressure which we can denote by $-p_F I$. Let F' be the deformation gradient of another configuration of the isotropic material. Then the classical infinitesimal strain tensor, E, measuring the "strain" on going from F to F', is given by the symmetric part of $F'F^{-1} - I$:

$$H = F'F^{-1} - I, \tag{4.6a}$$

$$E = \tfrac{1}{2}(H + H^T). \tag{4.6b}$$

If F' is "close to" F, then, to within terms of order one in the norm of H, $\mathcal{S}(F')$ is given by

$$\mathcal{S}(F') \simeq -p_F I + \lambda_F(\operatorname{tr} E)I + 2\mu_F E. \tag{4.7}$$

Here λ_F and μ_F, called *Lamé coefficients*, are scalars depending on F (and thus on p_F). In classical elasticity theory, it is shown that the quantity $\lambda_F + \frac{2}{3}\mu_F$ is the modulus of compressibility for the state F, i.e. $\lambda_F + \frac{2}{3}\mu_F$ is equal to our $k(F)$. Hence, for undistorted states F in isotropic materials, Proposition 2 yields

$$\lambda_F + \tfrac{2}{3}\mu_F \geq \tfrac{2}{3}p_F. \tag{4.8}$$

We believe that in physical applications this inequality should be found to hold for all isotropic elastic solids.

If the undistorted configuration described by F is stress-free, then $p_F = 0$, and (4.8) yields the classical inequality (1a) with, of course, $>$ weakened to $\geq$.

Fluids: An elastic fluid is a special isotropic material for which $S = -p(\rho)I$ for all F, and $k = \rho\ dp/d\rho$. Hence, for such a material, (4.4) is equivalent to (2b).

[7] Of course, a small uniform expansion will, in general, cause a change dT in T as well as a change $d\bar{p}$ in $\bar{p}$, but the change in T does not interest us here.

Study of Several Lattice Systems with Long-Range Forces

M. KAC
Rockefeller Institute, New York, New York
AND
E. HELFAND
Bell Telephone Laboratories, Incorporated, Murray Hill, New Jersey
(Received 5 April 1963)

A number of one- and two-dimensional Ising lattice systems with long-range ferromagnetic interactions are studied. The theory introduces as basic variables stochastic fields acting at each site, but goes beyond Weiss mean-field theory (or the Bragg–Williams approximation) in giving a complete account of the statistics of these fields. A transition is manifest in these systems by a shift in the values of the stochastic fields which are important for the calculation of the partition function. Particular attention is devoted to the critical region where the range of significant stochastic fields broadens. The equation of state for the lattice gas corresponding to this model is of the van der Waals type. Comparison is frequently made between these results and the properties of an analogous one-dimensional continuum system studied by Kac, Uhlenbeck, and Hemmer.

1. INTRODUCTION

RECENTLY, a method of evaluating the classical partition function has been suggested.[1] Basically, the partition function is expressed as an average over a set of stochastic fields acting on the particles as a result of intermolecular interactions. A general formulation of the procedure becomes tractable for the case of a one-dimensional system with exponential attraction, since then the stochastic processes involved are Markoffian. Explicitly, the theory has been developed for a one-dimensional continuum system of hard rods attracting each other with the potential $v(x) = -\alpha\gamma \exp(-\gamma|x|)$ and for a one-dimensional lattice model with pairwise site interaction $-J\gamma\mu_k\mu_{k'}$, $\exp(-\gamma|k-k'|)$, with $\mu = \pm 1$.

For a finite value of the interaction range parameter, γ, the one-dimensional continuum model exhibits no phase transition, as expected. In the limit as $\gamma \to 0$, Kac, Uhlenbeck, and Hemmer[2] have demonstrated that the equation of state in the one-phase region becomes the van der Waals equation. However, Maxwell tie lines appear explicitly in the two-phase region, rather than the van der Waals loops. In subsequent works, discussions of the pair correlation function[3] and of critical phenomena[4] have been presented.

The properties of the lattice systems to be studied in this paper are closely related to the above mentioned continuum system. An added degree of simplicitly is achieved due to the nearly trivial nature of the short-range force of the lattice model. Thereby a more concise form of the mathematical development is possible. In Sec. 2, the general structure of the theory will be reviewed. A somewhat detailed treatment of the one-dimensional field-free Ising model with infinite-range exponential attraction is presented in order to expound the essential features of such systems. A transition occurs at the Weiss–Bragg–Williams transition point. The transition is manifest by nonzero stochastic fields (although equally likely positive as negative) making important contributions to the partition function. The lattice gas with this interaction (corresponding to the ferromagnet in a field) again has a van-der-Waals-like equation of state. Of particular interest are some unusual features of the system in the region of the critical point. The pair correlation is also studied and found to be similar to the long-range part of the continuum gas correlation function.

The methods employed for the linear array lean heavily on the Markoffian nature of the probability process involved in the functional integration. Markoff processes are fundamentally one-dimensional, so the question arises as to how these techniques may be adopted for the treatment of two-dimensional problems. Such systems are of interest since phase transitions occur even for finite-range potentials. Several such planar problems are discussed in Secs. 6–9. These systems involve potentials which are products of two functions, one dependent on the row distance, and the other on the column distance. Mathematically, the treatment involves considera-

[1] M. Kac, Phys. Fluids 2, 8 (1959). A brief review and qualitative discussion is given in E. Helfand, Ann. Rev. Phys. Chem. 14, (1963) (to be published). Reference is there made to the related lines of development being pursued by A. J. F. Siegert and by G. A. Baker, Jr., as well as to various quantum statistical applications.
[2] M. Kac, G. E. Uhlenbeck, and P. C. Hemmer, J. Math. Phys. 4, 216 (1963).
[3] G. E. Uhlenbeck, P. C. Hemmer, and M. Kac, J. Math. Phys. 4, 229 (1963).
[4] P. C. Hemmer, M. Kac, and G. E. Uhlenbeck, J. Math. Phys. (to be published).

tion of a multidimensional functional integral; i.e., a stochastic field variable is introduced for each row. This leads to collective effects among the stochastic fields in different rows. The properties of the two-dimensional systems are not qualitatively different from the one-dimensional systems when the interaction is long range in at least one direction (cf. Baker[5]).

2. GENERAL THEORY

The configuration probability distribution function of classical statistical mechanics may be viewed as a multidimensional Gaussian function. For example, for the Ising problem with general pair interaction $\mu_j\mu_k v(\mathbf{jk})$ between the **j**th and **k**th site, the probability of the configuration $\{\mu\} = \mu_1, \mu_2, \cdots, \mu_N$, where $\mu_i = \pm 1$, is

$$P_{\{\mu\}} = \exp\left[-\tfrac{1}{2}\beta \sum_{j,k=1}^{N}{}' \mu_j\mu_k v(\mathbf{jk})\right] \Big/ Q_N, \tag{2.1}$$

$$Q_N = \sum_{\{\mu\}} \exp\left[-\tfrac{1}{2}\beta \sum_{j,k}^{N}{}' \mu_j\mu_k v(\mathbf{jk})\right] \tag{2.2}$$

(prime on $\sum$ denotes $j \neq k$). It is perhaps not surprising that mathematical techniques from the theory of probability, where Gaussian forms have long been of importance, are of value in the problem of evaluating the partition function. Thus, one has the mathematical identity

$$Q_N = \sum_{\{\mu\}} \exp\left[-\tfrac{1}{2}\beta \sum_{j,k=1}^{N}{}' \mu_j\mu_k v(\mathbf{jk})\right]$$

$$= E\left\{\sum_{\{\mu\}} \exp\left[\beta^{\frac{1}{2}} \sum_{k=1}^{N} X(\mathbf{k})\mu_{\mathbf{k}}\right]\right\} \exp\left[\tfrac{1}{2}N\beta v(0)\right], \tag{2.3}$$

where the functional integration,[6] or expectation value, $E\{\ \}$, is to be taken over the Gaussian random process $X(\mathbf{k})$, characterized by the mean

$$E\{X(\mathbf{k})\} = 0, \tag{2.4}$$

and covariance

$$E\{X(\mathbf{j})X(\mathbf{k})\} = -v(\mathbf{jk}), \tag{2.5}$$

provided, of course, that such a process exists.

The partition function, aside from the expectation operation, has a form similar to that for a set of independent particles in an external field $-\beta^{\frac{1}{2}}X(\mathbf{k})$. In this sense the theory is much like Weiss mean-field theory (or Bragg–Williams theory), except that now with proper account of the stochastic field, exact results may be retrieved. The summation over $\{\mu\}$ may be performed for this set of "independent" particles:

$$Q_N = E\left\{\prod_{k=1}^{N} \cosh \beta^{\frac{1}{2}}X(\mathbf{k})\right\} 2^N \exp\left[\tfrac{1}{2}N\beta v(0)\right]. \tag{2.6}$$

All the difficulty now resides in performing the functional integration.

There are several routes along which one may proceed. Various iteration procedures lead to the equivalent of diagram expansions. Rather than pursue this tack, we will continue along lines previously employed,[1] and consider a class of problems such that the Gaussian random process $X(\mathbf{k})$ is Markoffian. This is the case if the system is one-dimensional and the intermolecular potential is an exponential;

$$v(kk') = -J\gamma \exp\left(-\gamma\,|k - k'|\right). \tag{2.7}$$

In this case, the X process, with covariance

$$E\{X(k)X(k')\} = \exp\left(-\gamma\,|k - k'|\right) \tag{2.8}$$

(the constants of the potential are conveniently absorbed into X), is an Ornstein–Uhlenbeck process,[7] first studied in connection with the theory of Brownian motion. As has been shown,[1] the evaluation of the expectation value involved in the partition function is equivalent to the determination of the highest eigenvalue of an operator. This becomes apparent upon explicitly writing

$$E\left\{\prod_{k=1}^{N} \cosh\left[(\nu\gamma)^{\frac{1}{2}}X(k)\right]\right\}$$

$$\equiv \int_{-\infty}^{\infty} \cdots \int \prod_{k=1}^{N} \cosh\left[(\nu\gamma)^{\frac{1}{2}}x_k\right]$$

$$\times W(x_1)P(x_1 \mid x_2, \gamma)P(x_2 \mid x_3, \gamma)$$

$$\cdots P(x_{N-1} \mid x_N, \gamma)\, dx_1 \cdots dx_N, \tag{2.9}$$

where a reduced inverse temperature $\nu = \beta J$ has been defined, and the Markoffian joint distribution has been written as a chain of pair distributions. The singlet probability distribution is

$$W(x) = (2\pi)^{-\frac{1}{2}} \exp\left(-\tfrac{1}{2}x^2\right), \tag{2.10}$$

and the pair distribution is

$$P(x \mid x', \gamma) = \left[2\pi(1 - e^{-2\gamma})\right]^{-\frac{1}{2}}$$

$$\times \exp\left[-(x' - xe^{-\gamma})^2/2(1 - e^{-2\gamma})\right]. \tag{2.11}$$

Chain integrals of the type appearing in Eq. (2.9)

[5] G. A. Baker, Jr., Phys. Rev. **130**, 1406 (1963).

[6] I. M. Gel'fand and A. M. Yaglom, Usp. Mat. Nauk **11**, 77 (1956), [English translation: J. Math. Phys. **1**, 48 (1960)]; S. G. Brush, Rev. Mod. Phys. **33**, 79 (1961); M. Kac, *Probability and Related Topics in Physical Sciences* (Interscience Publishers, Inc., New York, 1959), Chap. IV.

[7] G. E. Uhlenbeck and L. S. Orstein, Phys. Rev. **36**, 823 (1930).

may be viewed as operator multiplications in coordinate representation. Thus, consider an integral operator with kernel

$$K(x \mid x', \gamma) = \cosh^{\frac{1}{2}} [(\nu\gamma)^{\frac{1}{2}}x]S(x \mid x', \gamma) \cosh^{\frac{1}{2}} [(\nu\gamma)^{\frac{1}{2}}x'], \tag{2.12}$$

$$S(x \mid x', \gamma) = W(x)P(x \mid x', \gamma)/[W(x)W(x')]^{\frac{1}{2}}. \tag{2.13}$$

The eigenvalues of the Kac equation

$$\int_{-\infty}^{\infty} K(x \mid x', \gamma)\varphi(x')\, dx' = \lambda\varphi(x), \tag{2.14}$$

which we label $\lambda_0 \geq \lambda_1 \geq \lambda_2, \cdots$, correspond to orthonormal eigenfunctions $\varphi_0, \varphi_1, \cdots$. It has been previously shown that these eigenfunctions may be employed as a basis allowing one to write for the partition function

$$Q_N = \sum_{i=0}^{\infty} \lambda_i^{N-1} A_i^2 (2e^{-\frac{1}{2}\nu\gamma})^N, \tag{2.15}$$

where

$$A_i = \int_{-\infty}^{\infty} W^{\frac{1}{2}}(x) \cosh^{\frac{1}{2}} [(\nu\gamma)^{\frac{1}{2}}x]\varphi_i(x)\, dx. \tag{2.16}$$

In the thermodynamic limit of $N \to \infty$ only the largest eigenvalue contributes to the sum (2.15) and we obtain as an expression for the partition function

$$\lim_{N\to\infty} Q_N^{1/N} = 2e^{-\frac{1}{2}\nu\gamma}\lambda_0. \tag{2.17}$$

3. ONE-DIMENSIONAL FERROMAGNET

Recalling the relation of the Ornstein–Uhlenbeck process to the Fokker–Planck equation,[7] one may easily show that $S(x \mid y, \gamma)$ is the fundamental solution to the equation

$$\{\partial/\partial\gamma - [(\partial^2/\partial x^2) - \tfrac{1}{4}x^2 + \tfrac{1}{2}]\} S(x \mid y, \gamma) = 0, \tag{3.1}$$

subject to the initial condition

$$S(x \mid y, 0) = \delta(x - y). \tag{3.2}$$

Therefore S may be formally written as

$$S(x \mid y, \gamma) = \exp \{\gamma[(\partial^2/\partial x^2) - \tfrac{1}{4}x^2 + \tfrac{1}{2}]\} \delta(x - y). \tag{3.3}$$

Representation of the kernel in terms of the Dirac δ function will be of particular value in simplifying the lattice problems considered in this paper. This representation does not appear to be of comparable value in the continuum problems previously studied.

By employing the δ function, one may write the Kac equation (2.14) as

$$\exp \{\tfrac{1}{2} \log \cosh [(\nu\gamma)^{\frac{1}{2}}x]\} \exp \{\gamma[(d^2/dx^2) - \tfrac{1}{4}x^2]\} \times \exp \{\tfrac{1}{2} \log \cosh [(\nu\gamma)^{\frac{1}{2}}x]\}\varphi(x) = \lambda e^{-\frac{1}{2}\gamma}\varphi(x). \tag{3.4}$$

If the exponentials could be combined into a single operator, the problem would be considerably simplified. Series rules for this combination will prove to be useful for the long-range force, i.e., $\gamma \to 0$ limit, wherein our interests lie.

The basic theorems for the combination of exponentials of operators are due to Baker and Hausdorff.[8] They enable us to write

$$e^A e^B = e^C, \tag{3.5}$$

with

$$C = A + B + \tfrac{1}{2}[A, B] + \tfrac{1}{12}[A, [A, B]] - \tfrac{1}{12}[B, [A, B]] - \tfrac{1}{24}[A, [B, [A, B]]] + \text{commutators involving 5 or more } A\text{'s and } B\text{'s}. \tag{3.6}$$

The Baker–Hausdorff method of determining higher terms is iterative, while the form resulting from Lie algebra[9] is not very compact. For our purposes, we are interested in the variant

$$e^{\frac{1}{2}B}e^A e^{\frac{1}{2}B} = \exp \{A + B + \tfrac{1}{12}[A, [A, B]] + \tfrac{1}{24}[B, [A, B]] + \cdots\}. \tag{3.7}$$

For a general γ, each term of the series on the right-hand side might be important and this approach may not be the most expedient. For small γ, however, the operators $\gamma[(d^2/dx^2) - \frac{1}{4}x^2]$ and $\log \cosh [(\nu\gamma)^{\frac{1}{2}}x]$ may be regarded as commuting since

$$[\gamma\{(d^2/dx^2) - \tfrac{1}{4}x^2\}, \log \cosh \{(\nu\gamma)^{\frac{1}{2}}x\}] = -2\gamma^{\frac{3}{2}}\nu^{\frac{1}{2}} \tanh [(\nu\gamma)^{\frac{1}{2}}x]\, d/dx - \nu\gamma^2 \operatorname{sech}^2 [(\nu\gamma)^{\frac{1}{2}}x]. \tag{3.8}$$

For phenomena involving x and variations of x of $O(1)$ this term is of $O(\gamma^2)$, which is smaller than the $O(\gamma)$ contributions of the individual operators. Other cases will be considered below, but for all of them, this commutator, and to an even greater extent, higher-order commutators, are negligible. Interest may therefore be centered on the eigenvalue problem

$$\{-(d^2/dx^2) + \tfrac{1}{4}x^2 - \gamma^{-1} \log \cosh [(\nu\gamma)^{\frac{1}{2}}x]\}\varphi(x) = \kappa\varphi(x), \tag{3.9}$$

$$\lambda_0 = \exp [\gamma(\tfrac{1}{2} - \kappa_0)], \tag{3.10}$$

where κ_0 is the lowest eigenvalue. The analogy to

[8] G. H. Weiss and A. A. Maradudin, J. Math. Phys. **3**, 771 (1962).

[9] N. Jacobson, *Lie Algebra* (Interscience Publishers, Inc., New York, 1962).

the problem of determining the ground-state energy of a particle in the external potential

$$U(x) = \tfrac{1}{4}x^2 - \gamma^{-1} \log \cosh [(\nu\gamma)^{\frac{1}{2}}x] \qquad (3.11)$$

will enable us to use the familiar concepts of quantum mechanics in discussing the solution of Eq. (3.9). For sufficiently low ν (high temperature), the log cosh function may be expanded to

$$\log \cosh [(\nu\gamma)^{\frac{1}{2}}x] = \tfrac{1}{2}\nu\gamma x^2 - \tfrac{1}{12}\nu^2\gamma^2 x^4, \qquad (3.12)$$

and to lowest order in γ, the equation is (the harmonic oscillator)

$$\{-(d^2/dx^2) + \tfrac{1}{4}(1 - 2\nu)x^2\}\varphi = \kappa\varphi, \qquad (3.13)$$

with eigenvalues

$$\kappa_j = (1 - 2\nu)^{\frac{1}{2}}(j + \tfrac{1}{2}), \quad j = 0, 1, 2, \cdots, \qquad (3.14)$$

and eigenfunctions, the Weber functions,

$$\varphi_j = [(1 - 2\nu)^{\frac{1}{2}}/2^n\pi^{\frac{1}{2}}n!]^{\frac{1}{2}} \times \exp [-\tfrac{1}{4}(1 - 2\nu)^{\frac{1}{2}}x^2]H_j[(1 - 2\nu)^{\frac{1}{4}}x], \qquad (3.15)$$

where H_j is the jth Hermite polynomial.

Taking $j = 0$ for the smallest κ, we obtain the partition function

$$\lim_{N\to\infty} \frac{1}{N} \log Q_N = \log 2 + \tfrac{1}{2}\gamma[(1 - \nu) - (1 - 2\nu)^{\frac{1}{2}}] + \cdots. \qquad (3.16)$$

The lowest-order term, log 2, is the entropy term obtained for a system of free spins; thus, for temperatures above any transition, the properties of the system become ideal as γ vanishes.

It is clear that some modification of the procedure must be made for $\nu > \frac{1}{2}$, since the coefficient of x^2 then becomes negative and the oscillator would undergo unbounded motion, at which point the MacLaurin expansion of log cosh is not valid. Equation (3.9) always has bound states, since for large $|x|$, the term γ^{-1} log cosh $[(\nu\gamma)^{\frac{1}{2}}x]$ behaves as $|x|$ and the $+\frac{1}{4}x^2$ dominates. The situation may be viewed as follows: At the origin there is a competition between $+\frac{1}{4}x^2$ and the $-\frac{1}{2}\nu x^2$ of the log cosh term to determine whether $U(x)$ starts out with positive or negative curvature. For $\nu > \frac{1}{2}$, the U function starts with a parabolic well in which the ground state is contained; i.e., the function φ_0 does not spread out far enough to feel the higher-order terms of log cosh. On the other hand, for $\nu < \frac{1}{2}$, the U curve turns down at the origin, but since it rises as $\frac{1}{4}x^2$ eventually, there must be two minima. The extrema of U are given by

$$U'(x_e) = \tfrac{1}{2}x_e - \nu^{\frac{1}{2}}\gamma^{-\frac{1}{2}} \tanh [(\nu\gamma)^{\frac{1}{2}}x_e] = 0. \qquad (3.17)$$

To keep contact with the notation of Kac, Uhlenbeck, and Hemmer,[2] let $x_e = 2^{\frac{1}{2}}\gamma^{-\frac{1}{4}}\eta$. Condition (3.17) becomes

$$\tanh [(2\nu)^{\frac{1}{4}}\eta] = (2\nu)^{-\frac{1}{4}}\eta. \qquad (3.18)$$

One solution is always $\eta = 0$, but examining

$$U''(2^{\frac{1}{4}}\eta/\gamma^{\frac{1}{2}}) = \tfrac{1}{2} - \nu \operatorname{sech}^2 [(2\nu)^{\frac{1}{4}}\eta], \qquad (3.19)$$

it is evident that for $\nu > \frac{1}{2}$, zero corresponds to a maximum. At $\nu = \frac{1}{2}$, two nonzero solutions of equal magnitude and opposite sign appear for condition (3.18). These represent the absolute minima. For $(1 - 2\nu) = 0(1)$, these minima are of distance $O(\gamma^{-\frac{1}{4}})$ from the origin and are of depth $O(\gamma^{-1})$. Therefore, the eigenfunctions may be constructed out of functions contained in each well. Let us expand U about the minimum, introducing the variable $w = x - \eta(2^{\frac{1}{4}}/\gamma^{\frac{1}{2}})$:

$$U(w + 2^{\frac{1}{4}}\gamma^{-\frac{1}{4}}\eta) = \gamma^{-1} \log \cosh [(2\nu)^{\frac{1}{4}}\eta] - \tfrac{1}{2}\gamma^{-1}\eta^2 - \tfrac{1}{4}w^2\{1 - 2\nu \operatorname{sech}^2 [(2\nu)^{\frac{1}{4}}\eta]\} + \cdots, \qquad (3.20)$$

where the next-order anharmonic term is of order $\gamma^{\frac{1}{2}}$. The validity of neglecting the commutator (3.8) should also be confirmed. For $w = O(1)$, the commutator is $O(\gamma^{\frac{3}{2}})$ compared to the parts of the operator up to $O(\gamma)$ which are retained. Therefore for temperatures below the critical, given by $\nu_c = \frac{1}{2}$, the eigenvalue equation is

$$[-(d^2/dw^2) + \tfrac{1}{4}w^2\{1 - 2\nu \operatorname{sech}^2 [(2\nu)^{\frac{1}{4}}\eta]\}]\varphi = \{\kappa - \gamma^{-1} \log \cosh [(2\nu)^{\frac{1}{4}}\eta] + \tfrac{1}{2}\gamma^{-1}\eta^2\}\varphi, \qquad (3.21)$$

which is identical in form to the previous Eq. (3.13), with

$$\nu \operatorname{sech}^2 [(2\nu)^{\frac{1}{4}}\eta] = \nu - \tfrac{1}{2}\eta^2 \quad \text{replacing} \quad \nu,$$

and

$$\kappa - \gamma^{-1} \log \cosh [(2\nu)^{\frac{1}{4}}\eta] + \tfrac{1}{2}\gamma^{-1}\eta^2 = \kappa + \tfrac{1}{2}\gamma^{-1} \log (1 - \tfrac{1}{2}\nu^{-1}\eta^2) + \tfrac{1}{2}\gamma^{-1}\eta^2 \quad \text{replacing} \quad \kappa.$$

Note that only even functions of η enter so that the result is independent of whether one takes the φ function centered about the positive or negative η, or a linear combination. We return to this point later.

The partition function is, therefore,

$$\lim_{N\to\infty} N^{-1} \ln Q_N = \ln 2 - \tfrac{1}{2}\eta^2 + \log \cosh [(2\nu)^{\frac{1}{4}}\eta] - \tfrac{1}{2}\gamma[1 - \nu + \{1 - 2\nu \operatorname{sech}^2 [(2\nu)^{\frac{1}{4}}\eta]\}^{\frac{1}{2}}] + \cdots,$$

$$= \ln 2 - \tfrac{1}{2}\eta^2 - \tfrac{1}{2} \log (1 - \tfrac{1}{2}\nu^{-1}\eta^2) - \tfrac{1}{2}\gamma[1 - \nu + (1 - 2\nu + \eta^2)^{\frac{1}{2}}] + \cdots; \qquad (3.22)$$

i.e., the long-range force has produced a collective effect on the partition function which is of $O(1)$. Note that this form for Q_N also covers temperatures above the transition point when one substitutes the appropriate $\eta = 0$. From the first equation of (3.22), it follows that the choice of η minimizes the free energy to $O(\gamma^0)$, as in Bragg–Williams theory.

The system is ideal above T_c. The transition is second-order with a finite discontinuity in the specific heat of $\frac{3}{2}k$ per particle. The $O(\gamma)$ term of the energy begins to grow as $\gamma(T - T_c)^{-\frac{1}{2}}$, as the critical point is approached, but as we shall see below, never actually diverges. The transition occurs at the temperature predicted by the Weiss mean-field (Bragg–Williams) theory,

$$1 = -(1/kT_c) \sum_{\substack{j=-\infty \\ \neq i}} v(ij) = 2J/kT_c. \qquad (3.23)$$

The region of the critical temperature deserves special attention, since at $\nu = \nu_c$ there is a vanishing of the quadratic term of the "potential" U which determines the range of the local stochastic fields which make important contributions to the partition function. The quartic term must then be retained. Qualitatively, the effects on the differential equation of having a flatter well is that larger values of x (larger stochastic fields) become important for the ground state. The results which follow are closely related to those which have been expounded for the continuum problem by Kac, Hemmer, and Uhlenbeck.[4]

With the quartic term of the expansion of log cosh, the eigenvalue equation (3.9) becomes

$$[-(d^2/dz^2) + \nu_1 z^2 + \tfrac{1}{12}z^4]\varphi = \kappa\gamma^{-\frac{1}{2}}\nu^{-\frac{1}{2}}\varphi, \qquad (3.24)$$

where $z = x\gamma^{-\frac{1}{2}}\nu^{-\frac{1}{2}}$, and the new temperature variable $\nu_1 = \frac{1}{4}\gamma^{-2/3}\nu^{-4/3}(1 - 2\nu)$ measures deviations from the critical temperature. For $(1 - 2\nu) = O(\gamma^{\frac{1}{2}})$, all the terms are of $O(1)$ when z variations are $O(1)$. Exactly at the critical temperature, κ is related to the lowest eigenvalue, σ_0, of the equation

$$[(d^2/dz^2) - \tfrac{1}{12}z^4]\varphi = \sigma\varphi. \qquad (3.25)$$

For temperatures just above the critical, $0 < \nu_1 \ll 1$, the quadratic term may be treated by perturbation theory. While a pair of minima do develop immediately below the transition temperature, they are not separated by a barrier of sufficient height or width to permit one to speak of a distinct eigenfunction in each region. Thus for $0 > \nu_1 \gg -1$, perturbation theory may still be employed. For $|\nu_1| = O(1)$, the full Eq. (3.24) must be studied. Apparently, it is necessary to speak of a transition temperature interval of $O(\gamma^{\frac{1}{2}})$. However, γ must vanish, and along with it this interval, in order to achieve any transition for the one-dimensional system.

We noted earlier that a term of the form $\gamma(T - T_c)^{-\frac{1}{2}}$ occurs in the energy as $T \to T_c$, but outside the $\gamma^{\frac{1}{2}}$ region. Thus, the largest this term could get before the considerations of the last few paragraphs must be used is $\gamma \cdot \gamma^{-\frac{1}{2}} = \gamma^{\frac{1}{2}}$. When the perturbation theory described above is employed, the energy goes smoothly from one to two phases.

4. THE LATTICE GAS

Rather than going on to a discussion of the properties of the ferromagnet in the presence of a magnetic field, let us study the corresponding properties of the lattice gas. This will enable us to make contact with the continuum gas results of Kac, Uhlenbeck, and Hemmer.[2]

Consider a system of N sites, each of which is either singly occupied ($\mu_i = +1$) or vacant ($\mu_i = -1$). The interaction between particles on sites i and j is again $-\alpha\gamma \exp[-\gamma|i - j|]$ so that the Hamiltonian is

$$H\{\mu\} = -\tfrac{1}{2}\alpha\gamma \sum_{j,k=1}^{N}{}' \tfrac{1}{2}(\mu_j + 1) \cdot \tfrac{1}{2}(\mu_k + 1) \cdot \exp[-\gamma|j - k|]. \qquad (4.1)$$

The grand partition function in the limit $N \to \infty$ is

$$G_N(z, T) = \sum_{\{\mu\}} z^{\sum \frac{1}{2}(\mu_j + 1)} \exp[-\beta H\{\mu\}]$$
$$= z^{\frac{1}{2}N} \exp\{\nu\gamma N[e^{-\gamma}/(1 - e^{-\gamma}) - \tfrac{1}{2}]\}$$
$$\times \sum_{\{\mu\}} \exp[\tfrac{1}{2}\nu \sum_{jk} \mu_j\mu_k e^{-\gamma|j-k|} + \zeta \sum_j \mu_j], \qquad (4.2)$$

where

$$\nu = \tfrac{1}{4}\beta\alpha \quad \text{and} \quad \zeta = \tfrac{1}{2}\ln z + [2\nu\gamma e^{-\gamma}/(1 - e^{-\gamma})].$$

(When ζ is interpreted as an external field divided by kT, this grand partition function is proportional to the partition function of the ferromagnet in field.) If the same stochastic process as was employed in the previous section is introduced, then

$$G_N = 2^N z^{\frac{1}{2}N} \exp\{\nu\gamma N[e^{-\gamma}/(1 - e^{-\gamma}) - \tfrac{1}{2}]\} \times E\{\prod_k \cosh[(\nu\gamma)^{\frac{1}{2}}X(k) + \zeta]\}; \qquad (4.3)$$

or, in the limit as $\gamma \to 0$,

$$G_N = 2^N z^{\frac{1}{2}N} \exp[\nu N - \tfrac{1}{2}\nu\gamma N + \tfrac{1}{2}\gamma N - \gamma\kappa_0 N], \qquad (4.4)$$

where κ_0 is the smallest eigenvalue of

$$[-(d^2/dx^2) + U(x, \zeta)]\varphi = \kappa\varphi \qquad (4.5)$$

$$U(x, \zeta) = \tfrac{1}{4}x^2 - \gamma^{-1} \log \cosh [(\nu\gamma)^{\frac{1}{2}}x + \zeta]. \quad (4.6)$$

The eigenfunctions center about the minima of U, $x_s = 2^{\frac{1}{2}}\gamma^{\frac{1}{2}}\eta$, where η is now given by

$$(2\nu)^{-\frac{1}{2}}\eta = \tanh [(2\nu)^{\frac{1}{2}}\eta + \zeta]. \quad (4.7)$$

At high temperatures only one solution exists. As the temperature is lowered, a second, but higher, well develops. This temperature represents not the phase transition, but the limit of metastability of the phase corresponding to the second well. For finite ζ the flattening of the U curve, such as lead to the critical region, occurs only where the metastable state is developing, and thus has no physical significance.

Next consider, for some temperature below the critical, the effect of varying ζ. For large ζ there is again only one solution to Eq. (4.7), i.e., one well of U. As ζ is lowered, a second, but higher, well develops. At $\zeta = 0$ the eigenvalues in the two wells are equal, and there are coexisting phases. As ζ is lowered further, the original phase now corresponds to the higher well and therefore can be thought of only in a metastable sense.

In terms of η, the grand partition function is quite close to the partition function of the field-free ferromagnet:

$$\lim_{N\to\infty} (1/N) \ln G_N = \beta p = \ln 2 + \tfrac{1}{2} \ln z + \nu + \tfrac{1}{2}\gamma(1 - \nu) - \tfrac{1}{2} \ln (1 - \eta^2/2\nu) - \tfrac{1}{2}\eta^2 - \tfrac{1}{2}\gamma(1 - 2\nu + \eta^2)^{\frac{1}{2}} + \cdots . \quad (4.8)$$

One may deduce from this the density in the one-phase region,

$$\rho = \tfrac{1}{2}(1 + (2\nu)^{-\frac{1}{2}}\eta) + O(\gamma), \quad (4.9)$$

or

$$z = e^{-8\nu\rho}(\rho/1 - \rho) + O(\gamma). \quad (4.10)$$

The equation of state in the one-phase region is

$$p + \alpha\rho^2 = p_s + O(\gamma). \quad (4.11)$$

where p_s is the equation of state of a gas interacting only with the short-range force; in this case,

$$\beta p_s = \ln [1/(1 - \rho)]. \quad (4.12)$$

Such van der Waals-like equations of state are to be expected for a long-range interaction, according to the derivation of van der Waals, arguments based on the virial theorem, or diagrammatic considerations.

In the two-phase region, the pressure is maintained constant as a result of the degeneracy of the Kac equation. The theory here so closely parallels that of the continuum gas,[2] that presentation again is unnecessary.

5. PAIR CORRELATION FUNCTION

The pair correlation function is a measure of the order imposed on the system as a result of the interactions. For those facts which we wish to develop here, let us return to the ferromagnetic model. The techniques outlined above are easily adopted to the calculation of

$$f(t) \equiv \langle \mu_l \mu_{l+t} \rangle \equiv Q_N^{-1} \sum_{\{\mu\}} \mu_l \mu_{l+t} \exp [\tfrac{1}{2}\nu\gamma \sum_{ij}{}' \mu_i \mu_j e^{-\gamma|i-j|}] = 2^N e^{-\frac{1}{2}N\nu\gamma} Q_N^{-1} E\Big\{\tanh [(\nu\gamma)^{\frac{1}{2}}X(l)] \tanh [(\nu\gamma)^{\frac{1}{2}}X(l + t)] \times \prod_{k=1}^{N} \cosh [(\nu\gamma)^{\frac{1}{2}}X(k)]\Big\}. \quad (5.1)$$

The expectation value may again be evaluated by going to a basis in which the kernel K is diagonal, viz., the eigenfunctions of the Kac equation (2.14). For notational simplicity, introduce the scalar product $(\varphi, \psi) = \int \varphi(x)\psi(x)\, dx$. We then have, for the expectation value in Eq. (5.1),

$$E\{\cdots\} = (W^{\frac{1}{2}}(x) \cosh^{\frac{1}{2}} [(\nu\gamma)^{\frac{1}{2}}x], \varphi_0)\lambda_0^{l-1} \times \Big[\sum_{i=0}^{\infty} (\varphi_0 \tanh [(\nu\gamma)^{\frac{1}{2}}x], \varphi_i)\lambda_i^t(\varphi_i, \tanh [(\nu\gamma)^{\frac{1}{2}}x]\varphi_0)\Big] \times \lambda_0^{N-l-t}(\varphi_0, \cosh^{\frac{1}{2}} [(\nu\gamma)^{\frac{1}{2}}x]W^{\frac{1}{2}}(x)). \quad (5.2)$$

Under the assumption that l and $l + t$ are far from the ends of the chain, only the highest eigenvalue, λ_0, has been retained in the first and last terms of the product. Because t is finite, however, all eigenvalues must be retained in the middle part of the product.

Inserting (5.2) and (2.17) for Q_N into (5.1) yields

$$f(t) = \sum_i (\lambda_i/\lambda_0)^t(\varphi_0, \tanh [(\nu\gamma)^{\frac{1}{2}}x]\varphi_i)^2. \quad (5.3)$$

A form more closely related to the continuum result[3] is obtained by considering the generating function [define $f(0) \equiv 1$]

$$F(z) = \sum_{t=0}^{\infty} f(t)z^t = \frac{\lambda_0}{z} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, \varphi_0(x) \tanh [(\nu\gamma)^{\frac{1}{2}}x] \times R(x, y; \lambda_0/z) \cdot \tanh [(\nu\gamma)^{\frac{1}{2}}y]\varphi_0(y), \quad (5.5)$$

where

$$R(x, y, s) = \sum_j \varphi_j(x)\varphi_j(y)/(s - \lambda_j) \qquad (5.6)$$

is the resolvent kernel of $K(x, y)$.

For the field-free ferromagnet, only the $j = 1$ term contributes to lowest order in γ. This yields, in the one-phase region,

$$f(t) = [\nu\gamma/(1 - 2\nu)^{\frac{1}{2}}] \exp [-\gamma t(1 - 2\nu)^{\frac{1}{2}}], \qquad (5.7)$$

which generalizes when a field is present, or for the lattice gas, to

$$\begin{aligned} f(t) &\equiv \langle\langle(\mu_i - \langle\mu\rangle)(\mu_j - \langle\mu\rangle)\rangle\rangle \\ &= \{\nu\gamma[1 - (\eta^2/2\nu)]^2/(1 - 2\nu + \eta^2)^{\frac{1}{2}}\} \\ &\quad \times \exp [-\gamma t(1 - 2\nu + \eta^2)^{\frac{1}{2}}]. \end{aligned} \qquad (5.8)$$

These results are identical, when appropriate translation is made, with the long-range part of the continuum fluid correlations [cf. Eq. (56) of reference 3].

In the two-phase region, the pair correlation is a linear combination, weighted by the volume fractions of each phase, of the correlation functions appropriate to each of the phases. This is based on precisely the same argument as was used for the continuum problem.[3]

As the critical temperature is approached, the more complicated eigenvalues and eigenfunctions of Eq. (3.24) [or (3.25) precisely at the critical point] must be employed. Again the $j = 1$ term is dominant but all odd j's contribute, even to $O(\gamma)$. Some idea of the magnitude may be of value. At $T = T_c$,

$$\begin{aligned} f(t) = \gamma^{2/3}[&0.656 \exp (-0.754\gamma^{4/3}t) \\ &+ 0.0019 \exp (-2.91\gamma^{4/3}t) \\ &+ 0.00006 \exp (-5.6\gamma^{4/3}t) + \cdots]. \end{aligned} \qquad (5.9)$$

6. TWO-DIMENSIONAL PROBLEMS

For one-dimensional systems with no singularities in the potential (other than the core for the continuum problems), it is necessary that the limit of infinite interaction range be taken in order to obtain a phase transition. On the other hand, two-dimensional systems already exhibit phase transitions with finite interaction ranges. Thus, it is of interest to consider a number of two-dimensional systems which may be analyzed in a fashion similar to that so far employed. The formalism at the beginning of Sec. 2 was not specific as to dimensions. Only when the system was chosen to yield a simple Markoff process was the specialization to one dimension made, since a Markoff process implies an ordering which can only be achieved in one dimension.

For a two-dimensional square lattice we shall now show that by introducing a set of Markoff processes, essentially one for each row, the techniques of the previous sections carry over. Again it will be necessary to postulate a special type of intersite interaction, viz., one which may be written as a product of a function of the horizontal separation and a function of the vertical separation. Systems of this category to be discussed are:

(A) A particle interacts with other particles in its row and the rows immediately above and below according to the exponential law

$$v(kl, k'l') = \begin{cases} -J\gamma e^{-\gamma|k-k'|}, & l' = l, \\ -\tau J\gamma e^{-\gamma|k-k'|}, & l' = l \pm 1, \\ 0, & \text{otherwise.} \end{cases} \qquad (6.1)$$

The system is periodic in the rows l.

(B) A particle interacts with other particles in its row exponentially, and with the particle immediately above and below it:

$$v(kl, k'l') = \begin{cases} -J\gamma e^{-\gamma|k-k'|}, & l' = l \\ -\tau J, & l' = l \pm 1, \quad k' = k, \\ 0, & \text{otherwise.} \end{cases} \qquad (6.2)$$

(C) A particle interacts with other particles with a potential which is an exponential of the distance between the two measured along the lattice bonds

$$v(kl, k'l') = -J\gamma^2 \exp [-\gamma(|k - k'| + |l - l'|)]. \qquad (6.3)$$

Note the necessity of introducing a prefactor of γ^2 to keep the total energy finite.

(D) Nearest-neighbor interactions, which shall be shown to be a special case of (A), (B), or (C).

Mathematically, model (A) is the simplest. We shall therefore devote our greatest attention to its study. The other interactions will be analyzed to the extent necessary to make clear the lines along which one should proceed, and to bring out characteristic properties. Since most of the salient features are already evident for the field free ferromagnet, we will concentrate mostly on this problem.

7. TWO-DIMENSIONAL SQUARE LATTICE SYSTEM WITH EXPONENTIAL INTERACTION ALONG A ROW AND THE NEAREST-NEIGHBOR ROWS

The partition function for the ferromagnet of N columns and M rows corresponding to interaction

potential A may be written in terms of a functional integral according to Eq. (2.6). The $\mathbf{k}$ of that equation is to be interpreted as a double index, k for column and l for row. One has

$$Q_{NM} = E\left\{\prod_{k=1}^{N}\prod_{l=1}^{M}\cosh\left[(\nu\gamma)^{\frac{1}{2}}X(k,l)\right]\right\}(2e^{-\nu\gamma/2})^{NM}, \quad (7.1)$$

where the mean of $X(k, l)$ is zero and the covariance is

$$E\{X(k,l)X(k',l')\} = e^{-\gamma|k-k'|}[\delta_{l',l} + \tau(\delta_{l',l-1} + \delta_{l',l+1})]. \quad (7.2)$$

A statistically equivalent description of such a stochastic process may be given in terms of M independent Ornstein–Uhlenbeck processes $U_l(k)$, with mean zero and covariance

$$E\{U_l(k)U_{l'}(k')\} = e^{-\gamma|k-k'|}\delta_{l',l}. \quad (7.3)$$

Define the constant ξ, related to the nearest-row interaction parameter τ of Eq. (6.1), by

$$\tau = [\xi(1-\xi)]^{\frac{1}{2}}. \quad (7.4)$$

It can easily be confirmed by direct substitution that equating the stochastic process $X(k, l)$ to

$$X(k,l) = \xi^{\frac{1}{2}}U_l(k) + (1-\xi)^{\frac{1}{2}}U_{l+1}(k) \quad (7.5)$$

yields the required covariance (7.2). In analogy to the one-dimensional procedure, the functional integral in expanded form is

$$E\left[\prod_{k=1}^{N}\prod_{l=1}^{M}\cosh\{(\nu\gamma)^{\frac{1}{2}}[\xi^{\frac{1}{2}}U_l(k) + (1-\xi)^{\frac{1}{2}}U_{l+1}(k)]\}\right]$$
$$= \int_{-\infty}^{\infty}\cdots\int\left[\prod_{k=1}^{N}C(\mathbf{u}_k)\right]W(\mathbf{u}_1)P(\mathbf{u}_1\mid\mathbf{u}_2,\gamma)$$
$$\cdots P(\mathbf{u}_{N-1}\mid\mathbf{u}_N,\gamma)\,d\mathbf{u}_1\cdots d\mathbf{u}_N, \quad (7.6)$$

where the set $\mathbf{u}_k \equiv u_{k1}, \cdots, u_{kl}, \cdots, u_{kM}$; $C(\mathbf{u}_k)$ is

$$C(\mathbf{u}_k) = \prod_{l=1}^{M}\cosh\{(\nu\gamma)^{\frac{1}{2}}[\xi^{\frac{1}{2}}u_{kl} + (1-\xi)^{\frac{1}{2}}u_{k,l+1}]\}, \quad (7.7)$$

and the probabilities $W(\mathbf{u})$ and $P(\mathbf{u}\mid\mathbf{u}',\gamma)$ are products of the corresponding one-dimensional probability functions, e.g.,

$$W(\mathbf{u}_k) = \prod_{l=1}^{M}W(u_{kl}), \quad (7.8)$$

$$P(\mathbf{u}_k\mid\mathbf{u}_{k'},\gamma) = \prod_{l=1}^{M}P(u_{kl}\mid u_{k'l},\gamma). \quad (7.9)$$

The partition function is related to the highest eigenvalue of an M dimensional Kac equation

$$\int_{-\infty}^{\infty}C^{\frac{1}{2}}(\mathbf{u})S(\mathbf{u}\mid\mathbf{u}',\gamma)C^{\frac{1}{2}}(\mathbf{u})\varphi(\mathbf{u}')\,d\mathbf{u} = \lambda\varphi(\mathbf{u}), \quad (7.10)$$

or equivalently,

$$\exp[\tfrac{1}{2}\log C(\mathbf{x})]\exp[\gamma(\nabla^2 - \tfrac{1}{4}\mathbf{x}\cdot\mathbf{x})]$$
$$\times\exp[\tfrac{1}{2}\log C(\mathbf{x})]\varphi(\mathbf{x}) = \lambda e^{-\frac{1}{2}M\gamma}\varphi(\mathbf{x}). \quad (7.11)$$

One has, as in one dimension,

$$\lim_{N,M\to\infty}Q_{NM}^{1/NM} = 2e^{-\frac{1}{2}\nu\gamma}\lim_{M\to\infty}\lambda_0^{1/M}. \quad (7.12)$$

To lowest order in γ, the exponentials may be combined so that one need study only the eigenvalue problem

$$\left\{\sum_{l=1}^{M}[-(\partial^2/\partial x_l)^2 + \tfrac{1}{4}x_l^2 - \gamma^{-1}\log\cosh((\nu\gamma)^{\frac{1}{2}}[\xi^{\frac{1}{2}}x_l + (1-\xi)^{\frac{1}{2}}x_{l+1}])]\right\}\varphi(\mathbf{x}) = \kappa\varphi(\mathbf{x}), \quad (7.13)$$

with

$$\lambda_0 = \exp(\tfrac{1}{2}M\gamma - \kappa_0\gamma).$$

For high enough temperatures, the log cosh may be expanded, and the equation becomes

$$\sum_l[-(\partial^2/\partial x_l^2) + \tfrac{1}{4}(1-2\nu)x_l^2 - \tfrac{1}{4}(2\nu\tau)(x_{l-1}x_l + x_lx_{l+1})]\varphi = \kappa\varphi, \quad (7.14)$$

or

$$[-\nabla^2 + \tfrac{1}{4}\mathsf{A}:\mathbf{xx}]\varphi = \kappa\varphi, \quad (7.15)$$

where A is the Toeplitz matrix (for an infinite system, $M\to\infty$) with $(1-2\nu)$ on the diagonal and $(-2\nu\tau)$ on the first off-diagonal positions. The partial differential equation is rendered separable by the transformation

$$y_n = M^{-\frac{1}{2}}\sum_{k=1}^{M}x_ke^{2\pi ikn/M} = y_{-n}^*, \quad (7.16)$$

with inverse

$$x_k = M^{-\frac{1}{2}}\sum_{n=-\Omega}^{\Omega}y_ne^{-2\pi ikn/M} = x_k^*. \quad (7.17)$$

For M odd, the index n takes on the M values from $-\frac{1}{2}(M-1)$ to $+\frac{1}{2}(M-1)$, while for M even, the limits are $-\frac{1}{2}(M-2)\le n\le\frac{1}{2}(M-1)$. Without loss of generality we assume M odd, and write for brevity, $\Omega = \frac{1}{2}(M-1)$.

With this transformation the eigenvalue equation becomes

$$\sum_{n=-\Omega}^{\Omega}[-(\partial^2/\partial y_n\,\partial y_{-n}) + \tfrac{1}{4}\omega_n^2y_ny_{-n}]\psi(\mathbf{y}) = \kappa\psi(\mathbf{y}), \quad (7.18)$$

where

$$\omega_n^2 = \sum_{j=-\infty}^{\infty} A_{s,s+j} e^{-2\pi i j n/M},$$
$$= (1 - 2\nu) - 4\nu\tau \cos(2\pi n/M),$$
$$= [1 - 2\nu(1 + 2\tau)] + 4\nu\tau[1 - \cos(2\pi n/M)]. \quad (7.19)$$

A complete separation of the partial differential equation is achieved in terms of the real and imaginary parts of y_n. Define

$$z_n = \sqrt{2}\,\mathrm{Re}\, y_n = \sqrt{2}\,\mathrm{Re}\, y_{-n}, \qquad z_n' = \sqrt{2}\,\mathrm{Im}\, y_n = -\sqrt{2}\,\mathrm{Im}\, y_{-n}; \quad (7.20)$$

then, grouping the $+n$ and $-n$ terms,

$$2\left(-\frac{\partial^2}{\partial y_n\,\partial y_{-n}} + \tfrac{1}{4}\omega_n^2 y_n y_{-n}\right) = \left[-\frac{\partial^2}{\partial z_n^2} - \frac{\partial^2}{\partial z_n'^2} + \tfrac{1}{4}\omega_n^2(z_n^2 + z_n'^2)\right]. \quad (7.21)$$

The lowest eigenvalue κ_0 is given by

$$\kappa_0 = \frac{1}{2}\sum_{n=-\Omega}^{\Omega} \omega_n \quad (7.22)$$

$$\to (M/2\pi)\int_0^{\pi} [(1 - 2\nu) - 4\nu\tau\cos\theta]^{\frac{1}{2}}\, d\theta. \quad (7.23)$$

A transition occurs at $\nu(1 + 2\tau) = \frac{1}{2}$, again the mean field value, when the lowest ω, ω_0, becomes negative. This indicates that variations of y_0 are becoming large. Since each x_k has a bit of the zeroth mode in it, each of the x_k's becomes large. Both qualitatively and quantitatively the situation is similar to the one-dimensional case. Each x must now be expanded about a displaced center, $\eta(2^{\frac{1}{2}}\gamma^{-\frac{1}{2}})$. Siegert[10] has investigated the question of whether a lower or equivalent minimum might be achieved by selecting different values, η_k, for each x_k, and found this not to be the case. The condition that η represent a minimum is

$$\tanh\{[2\nu(1 + 2\tau)]^{\frac{1}{2}}\eta\} = \eta[2\nu(1 + 2\tau)]^{-\frac{1}{2}}.$$

The new eigenvalue equation is identical with the old one, Eq. (7.11), with the replacement

$$\kappa' = \kappa - \gamma^{-1}\log\cosh\{[2\nu(1 + 2\tau)]^{\frac{1}{2}}\eta\} + \tfrac{1}{2}\gamma^{-1}\eta^2$$
$$= \kappa + \tfrac{1}{2}\gamma^{-1}\log[1 - \eta^2/2\nu(1 + 2\tau)] + \tfrac{1}{2}\gamma^{-1}\eta^2 \quad \text{replacing } \kappa, \quad (7.24)$$

$$\nu' = \nu\,\mathrm{sech}^2\{[2\nu(1 + 2\tau)]^{\frac{1}{2}}\eta\}$$
$$= \nu - \tfrac{1}{2}\eta^2/(1 + 2\tau) \quad \text{replacing } \nu. \quad (7.25)$$

The partition function both above and below the transition may be written in the unified fashion

$$\lim_{M,N\to\infty}(1/MN)\ln Q_{MN} = \ln 2 - \tfrac{1}{2}\eta^2 + \log\cosh\{[2\nu(1 + 2\tau)]^{\frac{1}{2}}\eta\} - \tfrac{1}{2}\gamma\Big\{1 - \nu + \pi^{-1} \times \int_0^{\pi}[(1 - 2\nu') - 4\nu'\tau\cos\theta]^{\frac{1}{2}}\, d\theta\Big\}, \quad (7.26)$$

since above the transition, $\eta = 0$ and $\nu' = \nu$. Comparing the result with the one-dimensional formula, one sees that, to zeroth order in γ, the only difference is the replacement of J with $J(1 + 2\tau)$. There is again a specific-heat discontinuity of $\frac{3}{2}k$, at the transition point. As the critical temperature is approached from above, a log $(T - T_c)$ growth occurs in the $O(\gamma)$ term of the energy.

The critical properties of the system are much more difficult to evaluate than in one dimension. The fourth-order term in the expansion of log cosh will involve mixtures of the normal variables, y_n. This problem is currently under investigation.

8. EXPONENTIAL INTERACTION ALONG A ROW AND NEAREST-NEIGHBOR INTERACTIONS IN THE COLUMNS

In order to study model B, defined by Eq. (6.2), write the random process $X(k, l)$ in terms of two sets of processes: the $U_l(k)$ of Eq. (7.3), with exponential covariance; and $Z_l(k)$, a set of independent, completely random, Gaussian processes, i.e., mean-zero, and covariance

$$E\{Z_l(k)Z_{l'}(k')\} = \delta_{k'k}\delta_{l'l}. \quad (8.1)$$

If we take

$$X(k, l) = \gamma^{\frac{1}{2}}U_l(k) + \xi^{\frac{1}{2}}Z_l(k) + (1 - \xi)^{\frac{1}{2}}Z_{l+1}(k), \quad (8.2)$$

then, as required,

$$E\{X(k, l)X(k', l')\} = \gamma e^{-\gamma|k-k'|}\delta_{l'l} + \delta_{kk'}\delta_{ll'} + \tau\delta_{k'k}[\delta_{l',l+1} + \delta_{l',l-1}]. \quad (8.3)$$

The partition function, as before, may be written as a functional integral, and is related [cf. Eq. (7.12)] to the highest eigenvalue of

$$\int V^{\frac{1}{2}}(\mathbf{u}, \mathbf{z})W^{\frac{1}{2}}(\mathbf{z})S(\mathbf{u} \mid \mathbf{u}', \gamma)W^{\frac{1}{2}}(\mathbf{z}')V^{\frac{1}{2}}(\mathbf{u}', \mathbf{z}') \times \Phi(\mathbf{u}', \mathbf{z}')\, d\mathbf{u}'\, d\mathbf{z}' = \lambda\Phi(\mathbf{u}, \mathbf{z}), \quad (8.4)$$

[10] A. J. F. Siegert (to be published).

where

$$V(\mathbf{u}, \mathbf{z}) = \prod_{l=1}^{M} \cosh \{\nu^{\frac{1}{2}}[\gamma^{\frac{1}{2}}u_l + \xi^{\frac{1}{2}}z_l + (1-\xi)^{\frac{1}{2}}z_{l+1}]\}. \tag{8.5}$$

As far as the z dependence is concerned, the kernel is in a factored form, so this variable is eliminated by multiplying Eq. (8.4) by $V^{\frac{1}{2}}(\mathbf{u}, \mathbf{z})W^{\frac{1}{2}}(\mathbf{z})$ and integrating over $\mathbf{z}$. The result is

$$\int D^{\frac{1}{2}}(\mathbf{u})S(\mathbf{u} \mid \mathbf{u}', \gamma)D^{\frac{1}{2}}(\mathbf{u}')\varphi(\mathbf{u}')\, du' = \lambda\varphi(\mathbf{u}), \tag{8.6}$$

where

$$D(\mathbf{u}) = \int V(\mathbf{u}, \mathbf{z})W(\mathbf{z})\, d\mathbf{z}, \tag{8.7}$$

and

$$\varphi(\mathbf{u}) = D^{-\frac{1}{2}}(\mathbf{u}) \int W^{\frac{1}{2}}(\mathbf{z})V^{\frac{1}{2}}(\mathbf{u}, \mathbf{z})\Phi(\mathbf{u}, \mathbf{z})\, d\mathbf{z}; \tag{8.8}$$

or in differential form

$$\exp [\tfrac{1}{2} \log D(\mathbf{x})] \exp [\gamma(\nabla^2 - \tfrac{1}{4}\mathbf{x}\cdot\mathbf{x})] \times \exp [\tfrac{1}{2} \log D(\mathbf{x})]\varphi(\mathbf{x}) = \lambda e^{-\frac{1}{2}M\gamma}\varphi(\mathbf{x}), \tag{8.8'}$$

with

$$\lim_{N,M\to\infty} Q_{NM}^{1/NM} = 2e^{-\frac{1}{2}\nu(\gamma+1)} \lim_{M\to\infty} \lambda_0^{1/M}. \tag{8.9}$$

For temperatures above any transition, D may be expanded for small γ. One finds, as $M \to \infty$,

$$D(\mathbf{x}) = D_0 \exp [\tfrac{1}{2}\nu\gamma \sum_{l,m} x_l x_m \tanh^{|l-m|} \nu\tau + \cdots], \tag{8.10}$$

where

$$D_0^{1/M} = e^{\frac{1}{2}\nu} \cosh \nu\tau$$

is related to the partition function when the columns do not interact.

To lowest order in γ,

$$\lambda_0 = D_0 \exp [\tfrac{1}{2}M\gamma - \kappa_0\gamma], \tag{8.11}$$

$$[-\nabla^2 + \tfrac{1}{2}\mathbf{B} : \mathbf{xx}]\varphi = \kappa\varphi, \tag{8.12}$$

$$B_{lm} = \begin{cases} 1 - 2\nu & l = m, \\ -2\nu t^{|l-m|} & l \neq m, \end{cases} \tag{8.13}$$

where $t = \tanh \nu\tau$.

A separation is again achieved with the variables y_n of the previous section. The ω's are given by

$$\omega_n^2 = (1 - 2\nu) - 4\nu \frac{t \cos (2\pi n/M) - t^2}{1 - 2t \cos (2\pi n/M) + t^2}. \tag{8.14}$$

When the smallest ω, namely, $\omega_0^2 = 1 - 2\nu e^{2\nu\tau}$, vanishes at ν_c given by

$$\nu_c e^{2\nu_c\tau} = \tfrac{1}{2}, \tag{8.15}$$

the solution becomes unstable. We must then expand $D(\mathbf{x})$ in the variables $w_k = x_k - \eta(2/\gamma)^{\frac{1}{2}}$ with η given by

$$\eta(2\nu)^{-\frac{1}{2}} = \sinh [(2\nu)^{\frac{1}{2}}\eta]/\{\sinh^2 [(2\nu)^{\frac{1}{2}}\eta] + e^{-4\nu\tau}\}^{\frac{1}{2}}, \tag{8.16}$$

so as to put one at the absolute minima.

The solution to lowest order in γ is then given by

$$\lambda_0 = \tilde{D}_0 \exp [-\tfrac{1}{2}M\eta^2 + \tfrac{1}{2}M\gamma - \kappa_0\gamma], \tag{8.17}$$

with

$$\tilde{D}_0^{1/M} = \tfrac{1}{2} \exp [\tfrac{1}{2}\nu(1 + 2\tau)]\{\cosh [(2\nu)^{\frac{1}{2}}\eta] + (\sinh^2 [(2\nu)^{\frac{1}{2}}\eta] + e^{-4\nu\tau})^{\frac{1}{2}}\}. \tag{8.18}$$

Equation (8.12) for κ_0 has $\mathbf{B}$ generalized by the replacement

$$\nu' = \nu - \tfrac{1}{2}\eta^2 \quad \text{replaces} \quad \nu, \tag{8.19}$$

and

$$t' = \frac{\cosh [(2\nu)^{\frac{1}{2}}\eta] - \{\sinh^2 [(2\nu)^{\frac{1}{2}}\eta] + e^{-4\nu\tau}\}^{\frac{1}{2}}}{\cosh [(2\nu)^{\frac{1}{2}}\eta] + \{\sinh^2 [(2\nu)^{\frac{1}{2}}\eta] + e^{-4\nu\tau}\}^{\frac{1}{2}}} \quad \text{replaces} \quad t. \tag{8.20}$$

This is essentially all that is needed to determine the partition function to $O(\gamma)$ both above and below the transition. The modification in the presence of an external field, or for the lattice gas, involves merely adding ζ onto $(2\nu)^{\frac{1}{2}}\eta$ on the right-hand side of Eq. (8.16) which defines η, just as Eq. (4.7) generalizes Eq. (3.18). These results will not be written out in detail here. Suffice it to mention that there is a specific-heat discontinuity of $6\nu_c k(1 + 2\nu_c\tau)(3 - 4\nu_c^2)^{-1}$ at the transition point, and that the $O(\gamma)$ term of the energy goes as $\gamma \log (T - T_c)$ as T_c is approached.

9. INTERACTION EXPONENTIAL IN THE "DISTANCE" MEASURED ALONG THE BONDS

To treat model C we will again introduce independent Ornstein–Uhlenbeck process $U_l(k)$ ($l = 0, 1, \cdots, M$), as defined by Eq. (7.3). The $X(k, l)$ process will have a covariance equal to

$$\exp [-\gamma(|k - k'| + |l - l'|)],$$

as required by our general theory, if we equate it statistically to

$$X(k, l) = U_0(k)e^{-\gamma l} + (1 - e^{-2\gamma})^{\frac{1}{2}} \sum_{p=1}^{l} U_p e^{-\gamma(l-p)}. \tag{9.1}$$

As usual, the properties of the system are related to a Kac equation, which in differential form is

$$\exp\left[\tfrac{1}{2}\log T(\mathbf{x})\right]\exp\left[\gamma(\nabla^2-\tfrac{1}{4}\mathbf{x}\cdot\mathbf{x})\right]\times\exp\left[\tfrac{1}{2}\log T(\mathbf{x})\right]\varphi(\mathbf{x})=\lambda e^{-\frac{1}{2}(M+1)\gamma}\varphi(\mathbf{x}),\quad(9.2)$$

where

$$T(\mathbf{x})=\prod_{l=1}^{M}\cosh\left\{\nu^{\frac{1}{2}}\gamma\left[x_0e^{-\gamma l}+(1-e^{-2\gamma})^{\frac{1}{2}}\sum_{p=1}^{l}x_pe^{-\gamma(l-p)}\right]\right\}.\quad(9.3)$$

For small γ, the exponentials may be combined, and we need study only

$$[-\nabla^2+\tfrac{1}{4}\mathbf{x}\cdot\mathbf{x}-\gamma^{-1}\log T(\mathbf{x})]\varphi=\kappa\varphi,\quad(9.4)$$

or, above the transition,

$$\left\{\left[-(\partial^2/\partial x_0^2)+\tfrac{1}{4}(1-\nu)x_0^2-2^{-\frac{1}{2}}\nu\gamma^{\frac{1}{2}}x_0\sum_{l=1}^{M}x_le^{-\gamma l}\right]+\left[\sum_{l=1}^{M}-(\partial^2/\partial x_l^2)+\tfrac{1}{4}\sum_{l,l'=1}(\delta_{ll'}-2\nu\gamma e^{-\gamma|l-l'|})x_lx_{l'}\right]\right\}\varphi=\kappa\varphi.\quad(9.5)$$

The x_0 variable has been separated out because its inclusion with the other variables would spoil the Toeplitz nature of the matrix of the quadratic form. The complete problem can be handled by the methods of bordered matrices,[11] but the single degree of freedom represented by x_0 makes contributions which relatively vanish as the number of other degrees of freedom becomes infinite. Dropping x_0, we can diagonalize Eq. (9.5) with the variables y_n of Sec. 7. The ω's are given by

$$\omega_n^2=1-4\nu\gamma^2\{2(1-\gamma)[1-\cos(2\pi n/M)]+\gamma^2[2-\cos(2\pi n/M)]+O(\gamma^3)\}^{-1}.\quad(9.6)$$

The lowest one is $n=0$,

$$\omega_0^2=1-4\nu,\quad(9.7)$$

so that a transition occurs at $\nu_c=\frac{1}{4}$.

The lowest eigenvalue κ_0 is given by

$$\kappa_0=\tfrac{1}{2}\sum_n\omega_n=\frac{M}{2\pi}\int_0^{\pi}\left[1-\frac{4\nu\gamma^2}{2(1-\gamma)(1-\cos\theta)+\gamma^2(2-\cos\theta)}\right]^{\frac{1}{2}}d\theta.\quad(9.8)$$

The second term of the radical is $O(\gamma^2)$ except for a region $|\theta|=O(\gamma)$ where it is $O(1)$. Thus the integral is[12]

$$\kappa_0=\tfrac{1}{2}M[1+O(\gamma)],\quad(9.9)$$

and the partition function is

$$\lim_{N,M\to\infty}(NM)^{-1}\ln Q_{MN}=\ln 2+O(\gamma^2).\quad(9.10)$$

It is not surprising that the first correction is $O(\gamma^2)$, in view of the fact that each pair interaction goes as γ^2. These terms arise not from a $O(\gamma^2)$ contribution from each normal mode (of the stochastic field on a particle in a row), but from a $O(\gamma)$ modification of $O(\gamma)$ of the modes (the long-wavelength ones). The first of these $O(\gamma)$'s is due to the long-range effects along a row, while the second arises from long-range effects between rows.

We could go on to study the region below the transition, or the lattice gas by shifting the center(s) of the eigenfunction(s) as before, where η is now given by

$$\eta/2\nu^{\frac{1}{2}}=\tanh(2\nu^{\frac{1}{2}}\eta+\zeta).\quad(9.11)$$

The result is that

$$\nu'=\nu-\tfrac{1}{4}\eta^2\quad\text{replaces}\quad\nu,$$

and

$$\kappa'=\kappa-\tfrac{1}{2}\gamma^{-1}\log(1-\tfrac{1}{4}\nu^{-1}\eta^2)-\tfrac{1}{2}\gamma^{-1}\eta^2\ \text{replaces}\ \kappa.$$

The outstanding thermodynamic properties are a specific-heat discontinuity of $\frac{3}{2}k$ at T_c, and the van der Waals equation of state

$$\beta p=-\ln(1-\rho)-8\nu\rho^2.\quad(9.12)$$

10. NEAREST-NEIGHBOR INTERACTIONS

The possibility of expressing the partition function in terms of the largest eigenvalue of an integral equation arises from the exponential nature of the potential function. It will now be demonstrated that nearest-neighbor interaction law is a limiting case of the exponential.

In one dimension, consider the potential

$$v(kk')=-Je^{\gamma}\exp(-\gamma|k-k'|).$$

In the limit as $\gamma\to\infty$, $v(kk')$ vanishes for $|k-k'|>1$, and is $-J$ for $|k-k'|=1$. As formulated above, the $k=k'$ term is retained in the functional integral. This may prove troublesome since $v(kk)$ diverges, but this self-energy is eventually subtracted out. By a modification of this type, models (A), (B), or (C), above, may be reduced to the two-dimensional nearest-neighbor problem.

[11] H. S. Wilf, *Mathematics for the Physical Sciences* (John Wiley & Sons, Inc., New York, 1962).

[12] *Note added in proof.* J. McKenna has shown that explicitly $\kappa_0=M\{1/2-\gamma(\nu/8\pi)\int_0^{\infty}Z^{-1/2}(1+Z)^{-1/2}[(1+Z)^{1/2}+(1-4\nu+Z)^{1/2}]^{-1}Z+\cdots\}$.

Irreducible Representations of Generalized Oscillator Operators*

T. F. JORDAN, N. MUKUNDA†, AND S. V. PEPPER
Department of Physics and Astronomy, University of Rochester, Rochester, New York
(Received 27 February 1963)

All of the irreducible representations are found for a single pair of creation and annihilation operators which together with the symmetric or antisymmetric number operator satisfy the generalized commutation relation characteristic of para-Bose or para-Fermi field quantization. The procedure is simply to identify certain combinations of these three operators with the three generators of the three-dimensional rotation group in the para-Fermi case, and with the three generators of the three-dimensional Lorentz group in the para-Bose case. The irreducible representations are then easily obtained by the usual raising and lowering operator techniques. The applicability of these techniques is demonstrated by a simple argument which shows that the commutation relations require that the generator to be diagonalized have a discrete spectrum.

I. INTRODUCTION

CONSIDERABLE recent attention has been given to generalized schemes of field quantization which would describe particles obeying statistics more general than the familiar Bose–Einstein or Fermi–Dirac types.[1-9] The essential new feature, from a mathematical point of view, is that the creation and annihilation operators obtained from the Fourier decomposition of the field are not required to satisfy the usual commutation or anticommutation relations. If we limit our attention to just a single pair of creation and annihilation operators $a^\dagger$ and a, the generalization of the commutation or anticommutation relations consists of the equation

$$[a, N] \equiv aN - Na = a, \tag{1}$$

together with an assumption that the number operator N has a particular form such as the symmetrized or antisymmetrized forms[3,7,9]

$$N = (\tfrac{1}{2})(a^\dagger a + aa^\dagger), \tag{B}$$

$$N = (\tfrac{1}{2})(a^\dagger a - aa^\dagger). \tag{F}$$

The purpose of this paper is to present a simple method for exhibiting all of the irreducible representations of operators $a^\dagger$ and a which satisfy Eq. (1), together with either (B) or (F). Since the familiar representations in which the operators $a^\dagger$ and a satisfy the usual commutation or anticommutation relations are particular solutions for the cases (B) and (F), respectively, we will refer to these cases as "para-Bose" and "para-Fermi".

For the para-Fermi case, our method consists simply of recognizing that the number operator, creation operator, and annihilation operator satisfy the same commutation relations as the third component and the raising and lowering operators for angular momentum.[10] The irreducible representations in this case correspond to the well known irreducible representations of the three-dimensional rotation group.

In the para-Bose case, the corresponding set of operators does not form a recognizable Lie algebra since it is not even closed under commutation. However, we do obtain a Lie algebra by considering the number operator together with the squares of the creation and annihilation operators. The appropriate combinations of these three operators satisfy commutation relations of generators of the three-dimensional (two space, one time) Lorentz group. From these commutation relations, plus the fact that the number operator (B) is positive, we can find the irreducible representations by the usual raising and lowering operator techniques. The irreducible representations of $a^\dagger$ and a are then obtained by taking square roots.

Our procedure is to find the irreducible representations of operators satisfying the *commutation*

* Supported in part by the United States Atomic Energy Commission.

† On leave of absence from the Atomic Energy Establishment, Bombay, India.

[1] G. Gentile, Nuovo Cimento **17**, 493 (1940).

[2] T. Okayama, Progr. Theoret. Phys. (Kyoto) **7**, 517 (1952).

[3] H. S. Green, Phys. Rev. **90**, 270 (1953).

[4] I. E. Mc Carthy, Proc. Cambridge Phil. Soc. **51**, 131 (1955).

[5] T. W. B. Kibble and J. C. Polkinghorne, Proc. Roy. Soc. (London) **A243**, 252 (1957).

[6] D. V. Volkov, Soviet Physics—JETP **9**, 1107 (1959) and **11**, 375 (1960).

[7] S. Kamefuchi and Y. Takahashi, Nucl. Phys. **36**, 177 (1962).

[8] G. F. Dell'Antonio, O. W. Greenberg, and E. C. G. Sudarshan, *Proceedings of the Istanbul Summer School (1962)*, edited by F. Gursey (Gordon and Breach, New York, 1963); see also University of Rochester Report NYO-10241 (unpublished).

[9] H. Scharfstein, Thesis (New York University, 1962).

[10] The connection between spin-1/2 angular-momentum operators, and creation and annihilation operators satisfying anticommutation relations is well known. P. Jordan and E. P. Wigner, Z. Phys. **47**, 631 (1928).

relations of the three-dimensional Lorentz group together with the requirement that the number operator be positive. If we were interested in finding actual irreducible representations of the Lorentz *group*, we would look for the irreducible representations of operators satisfying these same commutation relations but the positiveness would be replaced by a requirement insuring the compactness of the subgroup of (two-dimensional) space rotations. Thus, most of the representations that we find do not, in fact, generate representations of the three-dimensional Lorentz group, and many of the representations of generators of the three-dimensional Lorentz group, as found by Bargmann,[11] do not satisfy our requirement that the number operator be positive.

The raising and lowering operator techniques can be rigorously justified only if we know that the operator whose eigenvalues are being raised and lowered, namely the number operator, does in fact have a discrete spectrum. If we do not know this, we run the risk of overlooking representations in which this operator has a purely continuous spectrum. Hence we give a short proof that the commutation relations alone imply that the number operator has a discrete spectrum in each irreducible representation. This step is essential particularly in the para-Bose case; although the generator corresponding to the number operator is known to have a discrete spectrum for all irreducible representations of the three dimensional Lorentz group,[11] we find representations of the commutation relations which do not generate representations of the group. For the (compact) three-dimensional rotation group, it is of course true that representations of the commutation relations are equivalent to representations of the group. But we do not need to use either this fact or the general theorem that the irreducible unitary representations of a compact group are finite-dimensional and thus have generators with discrete spectra; we prefer to notice that the discreteness of the spectrum follows directly from the commutation relations.

The para-Fermi case is treated in Sec. II, and the para-Bose in Sec. III. In Sec. IV we make some remarks on the relation of our results to the work of Kamefuchi and Takahashi,[7] of O'Raifeartaigh and Ryan,[12] and of Greenberg and Messiah.[13]

[11] V. Bargmann, Ann. Math. **48**, 568 (1947).
[12] L. O'Raifeartaigh and C. Ryan, Proc. Roy. Irish. Acad. **A62**, 93 (1963).
[13] O. W. Greenberg and A. Messiah (to be published). We are indebted to Prof. Greenberg for communicating these results to us prior to publication.

II. PARA-FERMI OSCILLATOR

We want to find all of the irreducible representations of operators $a^\dagger$ and a which satisfy the commutation relation

$$[a, N] = a, \tag{1}$$

with the antisymmetric definition of the number operator

$$N = (\tfrac{1}{2})(a^\dagger a - aa^\dagger). \tag{F}$$

We define the three symmetric operators J_1, J_2, and J_3 by making the identifications

$$a^\dagger = J_1 + iJ_2, \qquad a = J_1 - iJ_2, \qquad N = J_3. \tag{2}$$

Equations (1) and (F) are then equivalent to the familiar angular-momentum commutation relations[10]

$$[J_i, J_j] = ie_{ijk}J_k, \qquad i, j, k = 1, 2, 3. \tag{3}$$

Our problem is equivalent to that of finding the irreducible representations of symmetric operators J_i which satisfy the commutation relations of generators of the three-dimensional rotation group.

Before referring to the well-known irreducible representations of angular momentum, we wish to remark that Eq. (3) implies that the operator J_3 has a discrete spectrum in each irreducible representation of J_1, J_2, and J_3. The usual treatment of angular momentum in terms of eigenvectors of J_3 is thus seen to be rigorously applicable. The proof is completely analogous to that given in Sec. III to establish the analogous property for the operator H_0. (Simply replace H_1, H_2, and H_0 everywhere in that proof by J_1, J_2, and J_3, respectively.)

By the usual raising and lowering operator techniques,[14] we can show that the familiar angular-momentum representations are all of the irreducible representations of symmetric operators J_1, J_2, and J_3, satisfying Eq. (3). All of the irreducible representations of operators $a^\dagger$ and a satisfying Eqs. (1) and (F) can be obtained by inverting Eqs. (2). We get a different irreducible representation for each different value

$$j = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, 2, \tfrac{5}{2} \cdots,$$

giving one of the possible eigenvalues $j(j+1)$ for the invariant operator

$$J^2 = J_1^2 + J_2^2 + J_3^2$$
$$= (\tfrac{1}{2})(a^\dagger a + aa^\dagger) + (\tfrac{1}{4})(a^\dagger a - aa^\dagger)^2.$$

The representation space in each case is $(2j+1)$-dimensional and the number operator has $(2j+1)$

[14] P. A. M. Dirac, *Quantum Mechanics* (Oxford University Press, London, 1958), pp. 144–149.

different eigenvalues. The first nontrivial representation is that for which $j = \frac{1}{2}$. In this particular case the representation is two-dimensional, the number operator has two different eigenvalues, and $a^\dagger$ and a are represented by the familiar 2×2 matrices which satisfy the anticommutation relation characteristic of Fermi–Dirac quantization.

III. PARA-BOSE OSCILLATOR

We want to find all of the irreducible representations of operators $a^\dagger$ and a which satisfy the commutation relation

$$[a, N] = a, \tag{1}$$

with the symmetric definition of the number operator

$$N = (\tfrac{1}{2})(a^\dagger a + a a^\dagger). \tag{B}$$

In this case, since N is not the commutator of $a^\dagger$ and a, these three operators do not form a Lie algebra. But the set of operators $a^\dagger a^\dagger$, aa, and N is closed under commutation. If we define the three symmetric operators H_1, H_2, and H_0 by

$$H_1 = (\tfrac{1}{4}i)(a^\dagger a^\dagger - aa), \qquad H_2 = (\tfrac{1}{4})(a^\dagger a^\dagger + aa),$$
$$H_0 = (\tfrac{1}{2})N,$$

we find in fact that the commutation relations

$$[H_2, H_0] = iH_1, \qquad [H_0, H_1] = iH_2, \qquad -[H_1, H_2] = iH_0 \tag{4}$$

follow from Eqs. (1) and (B).[15] These differ from the angular-momentum commutation relations (3) only by the presence of the minus sign in the last equation. They are the commutation relations satisfied by generators of the three-dimensional (two space, one time) Lorentz group. The operator H_0 corresponds to the generator of rotations of the two-dimensional space plane, and H_1 and H_2 correspond to the generators of Lorentz transformations in the two space directions.

Each irreducible representation of operators $a^\dagger$ and a satisfying Eqs. (1) and (B) will give symmetric operators H_1, H_2, and H_0 satisfying Eqs. (4). Our procedure is to find the relevant irreducible representations of H_1, H_2, and H_0, from which we will be able to work back to $a^\dagger$ and a. As a first step it is essential to show that H_0 has a discrete spectrum in each irreducible representation of symmetric operators H_1, H_2, H_0 satisfying Eqs. (4).

From Eqs. (4), it follows that the operators H_1 and H_2 transform as components of a vector under rotations generated by H_0;

$$e^{irH_0} H_1 e^{-irH_0} = H_1 \cos r - H_2 \sin r,$$
$$e^{irH_0} H_2 e^{-irH_0} = H_2 \cos r + H_1 \sin r.$$

By considering the case $r = 2\pi$, we find that the unitary operator $\exp(2\pi i H_0)$ commutes with each of the operators H_1, H_2, and H_0, so by Schur's lemma, $\exp(2\pi i H_0)$ must be a scalar multiple of the identity in each irreducible representation of H_1, H_2, and H_0. Since it is a unitary operator, we have in fact that

$$e^{2\pi i H_0} = e^{2\pi i d},$$

with d a real number. If x is any point in the spectrum of H_0, we thus have, by the spectral mapping theorem, that

$$e^{2\pi i x} = e^{2\pi i d},$$

or that

$$x = d + n,$$

with n an integer. We have shown that the spectrum of H_0 in each irreducible representation is discrete and consists of points separated by integer spaces. This property, which is strictly analogous to that of the angular-momentum operators, is, of course, not shared by the operators H_1 and H_2.

According to definition (B), the number operator is positive. Our next step is to find all of the irreducible representations of symmetric operators H_1, H_2, and H_0 satisfying Eqs. (4), with the additional restriction that H_0 be positive.

From Eqs. (4) it follows that the operator

$$Q = H_1^2 + H_2^2 - H_0^2$$

commutes with each of the operators H_1, H_2, and H_0, so by Schur's lemma, Q must be a scalar multiple of the identity in each irreducible representation of H_1, H_2, and H_0. Since Q is symmetric, we may assume that it is a real number. From the operator functions of H_1, H_2, and H_0, we may choose the complete set of commuting operators consisting of Q and H_0 (in analogy to J^2 and J_3 for angular momentum). The irreducible representations can be obtained by finding the simultaneous eigenvectors of Q and H_0 and constructing the matrices for H_1, H_2, and H_0 in this orthonormal basis. Since Q is simply a number in each irreducible representation, this amounts to finding the possible values for Q and the eigenvalues and eigenvectors of H_0 consistent with each value

[15] It has been pointed out to us by Dr. S. Okubo that this type of construction has also been used by H. J. Lipkin, "Collective Motion in Many-Particle Systems," *Brandeis University Summer Institute Lecture Notes* (W. A. Benjamin Company, Inc., New York, 1959).

of Q. The different values for Q will thus correspond to the different irreducible representations.

We define raising and lowering operators F and G by

$$F = H_1 + iH_2, \qquad G = H_1 - iH_2 = F^\dagger,$$

and find that

$$[H_0, F] = F, \tag{5}$$

$$[H_0, G] = -G, \tag{6}$$

$$F^\dagger F = GF = Q + H_0^2 + H_0, \tag{7}$$

$$G^\dagger G = FG = Q + H_0^2 - H_0. \tag{8}$$

Consider any irreducible representation of H_1, H_2, and H_0. We may assume that in this representation Q is some fixed real number, and H_0 is a positive operator with a discrete integrally spaced spectrum. Let $|h\rangle$ be an eigenvector of H_0 with eigenvalue h;

$$H_0 |h\rangle = h |h\rangle.$$

From Eq. (5) it follows that $F|h\rangle$ is either the zero vector or an eigenvector of H_0 with eigenvalue $h + 1$. If it is not zero, we will label it by its eigenvalue $h + 1$. In either case, assuming that all eigenvectors are to be normalized, it follows from Eq. (7) that

$$F |h\rangle = c(h)(Q + h^2 + h)^{\frac{1}{2}} |h + 1\rangle, \tag{9}$$

where $c(h)$ is a complex number with absolute value one. It follows similarly from Eq. (6) that $G|h\rangle$ is either the zero vector or an eigenvector of H_0 with eigenvalue $h - 1$. In either case, it follows from Eq. (8) that

$$G |h\rangle = c(h - 1)^*(Q + h^2 - h)^{\frac{1}{2}} |h - 1\rangle. \tag{10}$$

Starting with the given eigenvector of H_0 with eigenvalue h, we may repeatedly apply the operator G and construct eigenvectors of H_0 having eigenvalues successively decreased by one until we reach a vector on which G gives zero. Our requirement that H_0 be positive implies that we do in fact find such a vector—for otherwise, we would have negative eigenvalues for H_0. Let g be the eigenvalue of H_0 for this eigenvector on which G gives zero. Then from Eq. (8) or Eq. (10),

$$Q + g^2 - g = 0,$$

so that

$$Q = g - g^2 = g(1 - g). \tag{11}$$

The positiveness of H_0 requires that g be nonnegative. For the case $g = 0$, we have $Q = 0$, and it follows from Eq. (7) or Eq. (9) that F gives zero when applied to the eigenvector of H_0 with eigenvalue g. This vector thus yields the trivial one-dimensional representation with $H_1 = H_2 = H_0 = 0$. For a case where g is positive, we see from Eq. (7) or Eq. (9) that we can repeatedly apply the operator F, starting with the eigenvector of H_0 with eigenvalue g, and construct eigenvectors of H_0 having eigenvalues successively increased by one without ever reaching a vector on which F gives zero. The infinite sequence of eigenvectors thus constructed spans an infinite-dimensional space which is irreducible under the operation of F, G, and H_0, and therefore also irreducible under H_1, H_2, and H_0. We have shown that all of the irreducible representations with H_0 positive will in fact be found by this construction.

For each different value of g, which may be any nonnegative real number, we get a different irreducible representation of H_1, H_2, and H_0, satisfying Eqs. (4) with H_0 positive. Except for the trivial representation corresponding to $g = 0$, all of these representations are infinite-dimensional. The operator Q has the value $g(1 - g)$. The operator H_0 has the nondegenerate discrete spectrum consisting of the points $g + n$, with n a nonnegative integer. The corresponding eigenvectors $|g + n\rangle$ form an orthonormal basis in the representation space. In this basis, the matrix for H_0 is of course diagonal,

$$H_0 |g + n\rangle = (g + n) |g + n\rangle, \tag{12}$$

and the matrices for H_1 and H_2 can be obtained from those for F and G, which, from Eq. (9) and Eq. (10), are

$$\begin{aligned} F |g+n\rangle &= c(g+n)[g(1-g) + (g+n)^2 + g + n]^{\frac{1}{2}} \\ &\qquad \times |g + n + 1\rangle \\ &= c(g+n)[(2g+n)(n+1)]^{\frac{1}{2}} |g+n+1\rangle, \end{aligned} \tag{13}$$

$$\begin{aligned} G |g+n\rangle &= c(g+n-1)^* \\ &\qquad \times [g(1-g) + (g+n)^2 - g + n]^{\frac{1}{2}} \\ &\qquad \times |g + n - 1\rangle \\ &= c(g+n-1)^*[(2g+n+1)n]^{\frac{1}{2}} |g+n-1\rangle. \end{aligned} \tag{14}$$

As was mentioned in the introduction, most of the representations that we have found for symmetric operators H_1, H_2, and H_0 satisfying the commutation relations (4), are not suitable representations for generators of irreducible unitary representations of the three-dimensional Lorentz group. The reason is easy to see. If H_0 is actually to be the generator for rotations of the two-dimensional-space plane, it is necessary that

$$e^{2\pi i H_0} = \pm 1,$$

in order that rotations of 2π have the required character. This implies that

$$e^{2\pi i g} = \pm 1,$$

which restricts g to the values $g = (\frac{1}{2})m$, with m an integer. Only this subset of our representations are admissible as representation of the Lorentz group. On the other hand, many of the representations found by Bargmann[11] for generators of the Lorentz group are not of use to us because they fail to satisfy our additional requirement that H_0 be positive.

Let us now work backwards from the irreducible representations of H_1, H_2, and H_0 to construct the irreducible representations of $a^\dagger$, a, and N. Inversion of the equations defining H_1, H_2, H_0, F, and G gives

$$a^\dagger a^\dagger = -i2F, \tag{15}$$

$$aa = i2G, \tag{16}$$

$$N = 2H_0, \tag{17}$$

and Eqs. (12), (13), and (14) can be used to obtain the matrices for these operators in each irreducible representation of H_1, H_2, and H_0. It remains to take square roots.

From the commutation relation (1), we see that when the operators $a^\dagger$ and a operate on an eigenvector of N they give, if not the zero vector, an eigenvector of N with eigenvalue respectively increased or decreased by one. Each irreducible representation of $a^\dagger$ and a is defined on a vector space spanned by the nondegenerate eigenvectors of N. These eigenvectors correspond to integrally spaced eigenvalues which, since N is positive, range upwards from some lowest nonnegative eigenvalue. We distinguish two subspaces—the subspace spanned by eigenvectors corresponding to eigenvalues which differ from the lowest eigenvalue by an even nonnegative integer (even subspace), and the subspace spanned by eigenvectors corresponding to eigenvalues which differ from the lowest eigenvalue by an odd positive integer (odd subspace). These two subspaces together span the whole representation space for $a^\dagger$ and a. Our interest in them is due to the fact that the operators $a^\dagger a^\dagger$, aa, and N leave each of these subspaces invariant. Each irreducible representation of $a^\dagger$ and a gives rise to two irreducible representations of H_1, H_2, and H_0, one on the even subspace and the other on the odd subspace. In working backwards to get an irreducible representation of $a^\dagger$ and a, we must take the square root of the direct sum of two irreducible representations of H_1, H_2, and H_0.

Let us first combine the matrices for H_0 from two irreducible representations of H_1, H_2, and H_0 to obtain the matrix for N according to Eq. (17). Suppose that on the even subspace we choose the irreducible representation for H_1, H_2, and H_0, for which H_0 has the lowest eigenvalue g. Then $2g$ is the lowest eigenvalue of N, and N has the eigenvalues $2g + 2n$, with n a nonnegative integer, corresponding to the eigenvectors spanning the even subspace. Now we know that N must also have the eigenvalues $2g + 2n + 1$, so we have no choice but to choose H_0 on the odd subspace to have the eigenvalues $g + \frac{1}{2} + n$. Therefore, on the odd subspace we must choose the irreducible representation of H_1, H_2, and H_0, in which H_0 has the lowest eigenvalue $g + \frac{1}{2}$. If we order the eigenvectors of N (which form an orthonormal basis in the direct sum of the even and odd subspaces) according to the increasing eigenvalues of N, we effectively "fan together" the two irreducible representations, taking first a vector from one and then a vector from the other, and the matrix for N is

$$N = \begin{pmatrix} 2g & 0 & 0 & 0 & \\ 0 & 2g+1 & 0 & 0 & \cdots \\ 0 & 0 & 2g+2 & 0 & \\ 0 & 0 & 0 & 2g+3 & \\ & \vdots & & & \ddots \end{pmatrix},$$

with the first eigenvalue coming from one representation of H_0, the second from the other, etc.

Now we combine the same two irreducible representations of H_1, H_2, and H_0 on the even and odd subspaces, using Eqs. (13) and (15), to obtain the matrix for $a^\dagger a^\dagger$:

$$a^\dagger a^\dagger = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & \\ 2(2g)^{\frac{1}{2}} & 0 & 0 & 0 & 0 & \cdots \\ 0 & 2(2g+1)^{\frac{1}{2}} & 0 & 0 & 0 & \\ 0 & 0 & 2[(2g+1)2]^{\frac{1}{2}} & 0 & 0 & \\ 0 & 0 & 0 & 2[(2g+2)2]^{\frac{1}{2}} & 0 & \\ 0 & 0 & 0 & 0 & 2[(2g+2)3]^{\frac{1}{2}} & \\ & & \vdots & & & \ddots \end{pmatrix},$$

with the general matrix element being nonzero only in the cases

$$(a^\dagger a^\dagger)_{i+2,i} = [(4g + j - 1)(j + 1)]^{\frac{1}{2}} \quad \text{for } j \text{ odd},$$

$$(a^\dagger a^\dagger)_{i+2,i} = [(4g + j)j]^{\frac{1}{2}} \quad \text{for } j \text{ even}.$$

Here we have omitted both the phase factors c of Eq. (13) and the factor $-i$ of Eq. (15), it being understood that each matrix element of $a^\dagger a^\dagger$ is determined only to within a phase factor. The odd row–odd column matrix elements come from the irreducible representation of F, G, and H_0 on the even subspace, with H_0 having the lowest eigenvalue g, and the even row–even column matrix elements come from the irreducible representation of F, G, and H_0 on the odd subspace in which H_0 has the lowest eigenvalue $g + \frac{1}{2}$.

We know that $a^\dagger$ will have nonzero matrix elements only one place below the diagonal. We find that the only matrix of this kind which has a square equal to $a^\dagger a^\dagger$ and which gives the correct matrix for N is

$$a^\dagger = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ [2(2g)]^{\frac{1}{2}} & 0 & 0 & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & [2(2g+1)]^{\frac{1}{2}} & 0 & 0 & 0 \cdots \\ 0 & 0 & 0 & 4^{\frac{1}{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & [2(2g+2)]^{\frac{1}{2}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 6^{\frac{1}{2}} \\ & & \vdots & & & \ddots \end{pmatrix},$$

with the general matrix element being nonzero only in the cases

$$(a^\dagger)_{i+1,i} = (4g + j - 1)^{\frac{1}{2}} \quad \text{for } j \text{ odd},$$

$$(a^\dagger)_{i+1,i} = j^{\frac{1}{2}} \quad \text{for } j \text{ even},$$

which may also be written as

$$(a^\dagger)_{i+1,i} = [(2g - \tfrac{1}{2})(1 - (-1)^j) + j]^{\frac{1}{2}}.$$

It is to be understood that each matrix element of $a^\dagger$ is determined only to within a phase factor.

Each irreducible representation of operators $a^\dagger$ and a satisfying Eq. (1) with the symmetric definition (B) of the number operator N, is given by a matrix of the above form. We get a different non-trivial irreducible representation for each positive real value of g. (The irreducible representations of the operators H_1, H_2, and H_0, corresponding to $g = 0$, can only be combined to give the trivial one-dimensional representation with $a^\dagger = a = 0$.) For the particular case $g = \frac{1}{4}$, we get the familiar matrix

$$(a^\dagger)_{i+1,i} = j^{\frac{1}{2}},$$

which satisfies the commutation relation

$$[a, a^\dagger] = 1,$$

characteristic of Bose–Einstein quantization.[16]

[16] For the particular case of operators satisfying the usual commutation relations, the fact that there is just the one familiar irreducible representation was proved by J. von Neumann, Math. Ann. **104**, 570 (1931).

IV. DISCUSSION

An alternative approach to determining operators satisfying the commutation relation (1) with the symmetric or antisymmetric definition (B) or (F) of the number operator, is to find more restrictive equations satisfied by $a^\dagger$ and a such that the general commutation relation is automatically satisfied by the operators which satisfy the more restrictive equation. An example of such an equation is the familiar commutation relation or anticommutation relation characteristic of Bose–Einstein of Fermi–Dirac quantization. We have seen that of the many irreducible representations of operators $a^\dagger$ and a satisfying the general commutation relation (1) in the para-Bose or para-Fermi case there was just one representation which satisfied the familiar commutation or anticommutation relation, respectively.

A sequence of successively higher-order generalizations of the familiar commutation and anticommutation relations have been derived by Kamefuchi and Takahashi.[7] Each of these has the property of implying the general commutation relation (1) with the relevant definition (B) or (F) of the number operator. One can check, at least for the first few cases, that our irreducible representations of para-Fermi operators corresponding to $j = 1, \frac{3}{2}$, etc. satisfy the successively higher-order generalizations of the anticommutation relation (the representation for $j = \frac{1}{2}$ satisfies the ordinary anticommutation relations), and that our irreducible representations of para-Bose operators corresponding to $g = \frac{1}{2}, \frac{3}{4}$, etc. satisfy the successively higher-order generalizations of the commutation relations (for $g = \frac{1}{4}$, the representation satisfies the ordinary commutation relation). We evidently have one irreducible representation of operators $a^\dagger$ and a satisfying each of the higher-order generalizations of the commutation and anticommutation relations. We know of no such equation satisfied by our irreducible representations of para-Bose operators for other than quarter integral values of g.

The irreducible representations of the para-Bose operators $a^\dagger$ and a satisfying Eq. (1) with the definition (B) of the number operator, have also been found by O'Raifeartaigh and Ryan.[12] Besides providing an alternative way to arrive at these results, in what we think is a particularly simple and transparent manner, our method has the advantage of containing a proof that the number operator has the discrete spectrum necessary to make the eigenvector techniques applicable.

The representations of generalized oscillator operators have also been investigated by Greenberg and Messiah.[13] Their interest is in representations for which there is a unique vector that is annihilated by the operator a. They show that this vector is necessarily an eigenvector of the operator $aa^\dagger$. Let p be the corresponding eigenvalue. It is then shown that p is always an integer in the para-Fermi case, and in the para-Bose case it is argued that only integral values of p lead to reasonable results when several different sets of oscillator variables are combined to construct a quantized field. It is the representations having these particular properties which were established by Greenberg and Messiah. Let us determine which of our representations exhibit these properties.

For the para-Fermi case, consider any irreducible representation of $a^\dagger$ and a corresponding to some value of j. There exists in this representation a unique vector on which the operator a $(= J_1 - iJ_2)$ gives zero, namely the eigenvector of N $(= J_3)$ with lowest eigenvalue $-j$. This vector is an eigenvector of the operator

$$aa^\dagger = (J_1 - iJ_2)(J_1 + iJ_2) = J^2 - J_3^2 - J_3,$$

with eigenvalue

$$j(j+1) - j^2 + j = 2j.$$

Thus $p = 2j$ in always an integer. For each positive integer p, we have just one irreducible representation of $a^\dagger$ and a, and these are in fact all of the nontrivial irreducible representations.

For the para-Bose case, consider any irreducible representation of $a^\dagger$ and a corresponding to some value of g. We have again a unique vector on which the operator a gives zero, namely the eigenvector of N with lowest eigenvalue $2g$. By explicit matrix computation we find that this vector is an eigenvector of the operator $aa^\dagger$ with eigenvalue $4g$. Thus $p = 4g$ is an integer only for quarter-integral values of g which correspond exactly to the irreducible representations which satisfy the higher-order generalizations of the commutation relations derived by Kamefuchi and Takahashi.[7] For each positive integral value of p, we have just one irreducible representation of $a^\dagger$ and a, but in this case there are, of course, many other irreducible representations.

ACKNOWLEDGMENTS

We are indebted to Dr. A. J. Macfarlane for helpful comments and discussions, and are especially grateful to Professor E. C. G. Sudarshan for his continued interest in and helpful discussions of this work.

Formulation of the Many-Body Problem for Composite Particles*

M. GIRARDEAU†
The Enrico Fermi Institute for Nuclear Studies, The University of Chicago, Chicago, Illinois
(Received 7 March 1963)

The many-body problem for a system of composite particles is formulated in a way which takes explicit account of the composite nature of these particles and allows a clear separation between interatomic and intraatomic interactions. A second-quantization representation, fully equivalent to the conventional representation in which nuclei and electrons appear explicitly, is developed in terms of atomic annihilation and creation operators satisfying elementary Bose or Fermi commutation relations. All effects of the composite nature of the atoms are exactly contained in the interatomic and intraatomic matrix elements and in certain exchange integrals. An application is made to the problem of Bose condensation of fermion pairs.

1. INTRODUCTION

THERE are two distinct problems in treating a system of interacting particles each of which is composite in the sense of being composed of several (more) elementary constituents. The first is the well-known very difficult problem of finding reasonably accurate approximate solutions of any nontrivial quantum-mechanical many-body problem. The second problem, which would not occur if all particles present were elementary, is that of even formulating the problem in such a way as to take account of the existence of composite particles; this is the problem with which this paper is concerned. It is, of course, well known that composite particles behave like elementary bosons or fermions when they are (in some reasonable sense) well separated or when the interparticle interactions are small compared to the internal excitation energies,[1] but there are many problems in which these criteria are violated, yet the composite nature of the particles remains important; two examples are high-temperature gases and the electron-pair "quasibosons" of superconductivity theory. It would be very convenient if composite particles could also be treated as bosons or fermions in problems of this type. A different, but related, motivation is provided by the problem of the proper description of composite particles in relativistic quantum field theories. It is possible (in several ways) to define asymptotic field operators for composite particles satisfying local commutation relations, but local finite-time field operators for composite particles have not been defined. If this problem can be solved in the nonrelativistic case, the solution might admit a relativistic generalization.

The purpose of this paper is to show that a second-quantization representation for many-atom[2] systems can be developed in which the *atomic* annihilation and creation operators satisfy *elementary* boson or fermion commutation relations, i.e., the atoms behave like *point* particles. In this representation the Hamiltonian, expressed as a function of the local *atomic* field operators, takes the familiar form of a sum of a quadratic part representing independent-particle (here independent-atom) energies and a quartic part representing *two-body* interactions. The price one pays for this simplified formulation of the many-atom problem is that it is necessary to impose subsidiary conditions on the state space generated by the atomic creation operators in order that it be equivalent to the conventional-state space in which the atomic constituents (nuclei and electrons) are labelled explicitly; the physical-state vectors are required to be eigenstates of certain "exchange operators" in order that they correspond to vectors in the conventional-state space having the correct symmetry properties under exchange of nuclei or electrons between *different* atoms. These exchange operators have a simple explicit representation in terms of *atomic* annihilation and creation operators and numerical coefficients (exchange integrals). In this formalism, all effects of the composite nature of the atoms are exactly contained in the interatomic interaction matrix elements, which depend only upon the Coulomb interactions and the

* This work was supported in part by the United States Atomic Energy Commission.
† Present address: Institute of Theoretical Science and Department of Physics, University of Oregon, Eugene, Oregon.
[1] The formal proof was first given by P. Ehrenfest and J. R. Oppenheimer, Phys. Rev. 37, 333 (1931).
[2] Throughout this paper we shall use the term "atom" in a very general sense, to mean any structured nonrelativistic system composed of a small number of particles. Examples are real atoms or molecules and the electron-pair "quasibosons" of the theory of superconductivity. The terms "nucleus" and "electron" are, therefore, also to be interpreted in a general sense.

single-atom wavefunctions, and in the exchange integrals, which involve only the single-atom wavefunctions.

An elementary derivation of the atomic second-quantization formalism is carried out in Sec. 2. In Sec. 3 an alternative derivation is carried out by a method originally devised by Dyson[3] for the treatment of spin-wave interactions, in order to make contact with related work[3,4] and to clarify the significance of the exchange interactions. For purposes of illustration, the formalism is applied in Sec. 4 to systems of identical particles, by regarding *pairs* of particles as "quasiatoms." For the case of fermion pairs, the theory is compared with previous work of Blatt and Matsubara.[4] It is shown that the well-known results for the ideal Bose and Fermi gases are obtained with the new formalism involving *pair* annihilation and creation operators which satisfy Bose commutation relations regardless of whether the particles making up the pairs are bosons or fermions; Bose condensation of ideal fermion pairs, and more generally, multiple occupation of *single*-fermion states, is shown to be incompatible with the subsidiary condition. More generally, it is shown that even for a system of *interacting* fermions, *complete* Bose condensation of fermion pairs is incompatible with the exclusion principle for individual fermions; this has some bearing on the physical interpretation of the Bardeen–Cooper–Schrieffer theory of superconductivity.

2. ELEMENTARY DERIVATION OF ATOMIC SECOND-QUANTIZATION FORMALISM

Although our formalism is applicable to systems of arbitrary composite particles, we shall, for the sake of definiteness, consider in this section and the following one a system of identical atoms each containing one nucleus and l electrons. Let $\{\varphi_\alpha(Xx_1 \cdots x_l)\}$ be a set of single-atom wavefunctions, orthonormal and complete in the sense

$$\int \varphi_\alpha^*(Xx_1 \cdots x_l)\varphi_\beta(Xx_1 \cdots x_l)\, dX\, dx_1 \cdots dx_l = \delta_{\alpha\beta},$$

$$\sum_\alpha \varphi_\alpha^*(Xx_1 \cdots x_l)\varphi_\alpha(X'x_1' \cdots x_l') = (l!)^{-1}\delta(X - X') \times \sum_{P'} (-1)^{p(P')} P' \delta(x_1 - x_1') \cdots \delta(x_l - x_l'), \quad (1)$$

where $x_j = (\mathbf{r}_j\sigma_j)$ denotes both the position and spin-z component of electron j, X the position of the nucleus and also its spin-z component in case its total spin is zero, $\int$ means an integration over positions and summation over spins, $\delta_{\alpha\beta}$ is a Kronecker delta with respect to discrete and a Dirac delta function with respect to continuous quantum numbers, $\sum_\alpha$ is a sum over discrete and integral over continuous quantum numbers, and $\delta(X - X')$ and $\delta(x - x')$ are Dirac delta functions of position and Kronecker delta functions of spin. The form of the completeness relation takes into account the antisymmetry of the φ_α in the electron variables; $\sum_{P'}$ denotes a sum over all permutations P' of the primed variables, $p(P')$ being the parity of the permutation P'.

A system of nuclei and electrons whose numbers are appropriate to an integral number n of such atoms has a wavefunction ψ which can be expanded in terms of the single-atom wavefunctions as follows:

$$\psi(X_1 \cdots X_n x_1 \cdots x_{ln}) = \sum_{\alpha_1 \cdots \alpha_n} c(\alpha_1 \cdots \alpha_n) \times \varphi_{\alpha_1}(X_1 x_1 \cdots x_l) \cdots \varphi_{\alpha_n}(X_n x_{ln-l+1} \cdots x_{ln}), \quad (2)$$

with coefficients

$$c(\alpha_1 \cdots \alpha_n) = \int \varphi_{\alpha_1}^*(X_1 x_1 \cdots x_l) \cdots \varphi_{\alpha_n}^*(X_n x_{ln-l+1} \cdots x_{ln}) \times \psi(X_1 \cdots X_n x_1 \cdots x_{ln})\, dX_1 \cdots dX_n dx_1 \cdots dx_{ln}. \quad (3)$$

This does not in any way imply that ψ necessarily represents a state in which the nuclei and electrons are actually bound into atoms, since the unbound (continuum) wavefunctions are included in the complete set $\{\varphi_\alpha\}$.

It might be thought that the usefulness of an expansion such as (2) would be spoiled because we have picked one particular assignment of nuclei and electrons to atoms, i.e., nucleus 1 and electrons $1 \cdots l$ to atom 1, etc. This is, however, not the case; it follows from the antisymmetry of ψ in the electron coordinates, and its symmetry or antisymmetry in the nuclear coordinates, that either all of the coefficients c remain unchanged or else *all simultaneously* change sign under a permutation of the assignment of nuclei and electrons to atoms, depending on the parity of the permutation with respect to fermion variables. By similar reasoning, one can show that for a *given* assignment [namely, that in (3)], $c(\alpha_1 \cdots \alpha_n)$ is a symmetric or antisymmetric function of the α_i, depending upon whether $2J + l$ is even or odd, where J is the nuclear spin.

The function $c(\alpha_1 \cdots \alpha_n)$ may be thought of as a new wavefunction in which nuclear and electron variables no longer appear explicitly; every many-nucleus, many-electron wavefunction ψ de-

[3] F. J. Dyson, Phys. Rev. **102**, 1217 (1956).
[4] J. M. Blatt and T. Matsubara, Progr. Theoret. Phys. (Kyoto) **20**, 553 (1958), Sec. 4.

termines a many-atom wavefunction c according to (3). However, the converse is *not* true; there exist functions $c(\alpha_1 \cdots \alpha_n)$ whose correspondents $\psi(X_1 \cdots X_n\, x_1 \cdots x_{ln})$ have the wrong symmetry properties under exchange of nuclei and/or electrons between *different* atoms [with respect to the assignment in (2)], and hence cannot represent physical states of nuclei and electrons. It is therefore necessary to impose subsidiary conditions on the space of wavefunctions $c(\alpha_1 \cdots \alpha_n)$ in order that they all represent physically possible many-atom states. One requires in the many-nucleus, many-electron state space

$$\psi(X_1 \cdots X_q \cdots X_p \cdots X_n x_1 \cdots x_{ln}) = \pm\psi(X_1 \cdots X_p \cdots X_q \cdots X_n x_1 \cdots x_{ln}),$$

$$\psi(X_1 \cdots X_n x_1 \cdots x_{lp-l} x_{lq-l+1} x_{lp-l+2} \cdots x_{lq-l} x_{lp-l+1} x_{lq-l+2} \cdots x_{ln}) = -\psi(X_1 \cdots X_n x_1 \cdots x_{ln}), \quad (4)$$

where $1 \leq p < q \leq n$ and the plus or minus sign is to be taken in the top equation depending on whether the nuclei obey Bose or Fermi statistics. In the second of Eqs. (4), the electron variables exchanged are the farthest-left ones in atoms p and q; since [by (2)] ψ is automatically antisymmetric under electron exchanges *within* each atom, this entails no loss of generality. Substituting (2) and using the orthonormality of the φ_α, one finds that equivalent statements of (4), in terms of the many-atom wavefunction $c(\alpha_1 \cdots \alpha_n)$, are

$$\sum_{\alpha\beta} (\alpha_p\alpha_q\, |I_{\text{nuc}}|\, \alpha\beta) \times c(\alpha_1 \cdots \alpha_{p-1}\alpha\alpha_{p+1} \cdots \alpha_{q-1}\beta\alpha_{q+1} \cdots \alpha_n) = \pm c(\alpha_1 \cdots \alpha_n),$$

$$\sum_{\alpha\beta} (\alpha_p\alpha_q\, |I_{\text{elec}}|\, \alpha\beta) \times c(\alpha_1 \cdots \alpha_{p-1}\alpha\alpha_{p+1} \cdots \alpha_{q-1}\beta\alpha_{q+1} \cdots \alpha_n) = -c(\alpha_1 \cdots \alpha_n), \quad (5)$$

where the exchange operators I_{nuc} and I_{elec} for nuclei and electrons are defined in terms of their matrix elements, which are the nucleus and electron exchange integrals:

$$(\alpha_p\alpha_q\, |I_{\text{nuc}}|\, \alpha\beta) \equiv \int \varphi^*_{\alpha_p}(Xx_1 \cdots x_l)\varphi^*_{\alpha_q}(X'x_1' \cdots x_l') \times \varphi_\alpha(X'x_1 \cdots x_l)\varphi_\beta(Xx_1' \cdots x_l') \times dX\, dx_1 \cdots dx_l\, dX'\, dx_1' \cdots dx_l',$$

$$(\alpha_p\alpha_q\, |I_{\text{elec}}|\, \alpha\beta) \equiv \int \varphi^*_{\alpha_p}(Xx_1 \cdots x_l)\varphi^*_{\alpha_q}(X'x_1' \cdots x_l') \times \varphi_\alpha(Xx_1'x_2 \cdots x_l)\varphi_\beta(X'x_1x_2' \cdots x_l') \times dX\, dx_1 \cdots dx_l\, dX'\, dx_1' \cdots dx_l'. \quad (6)$$

The conditions (5) associated with different values of p and q are not independent; using the symmetry or antisymmetry of $c(\alpha_1 \cdots \alpha_n)$ in the α_i, it is easy to show that these conditions for one pair of values of p and q (e.g. $p = 1$, $q = 2$) imply the conditions for all other values of p and q. Rather than picking one particular pair of values, it is most convenient for the subsequent development to sum over all pairs (p, q) and hence to state the subsidiary conditions (5) in the symmetrized form

$$\sum_{\alpha\beta} \sum_{p<q}^{n} (\alpha_p\alpha_q\, |I_{\text{nuc}}|\, \alpha\beta) \times c(\alpha_1 \cdots \alpha_{p-1}\alpha\alpha_{p+1} \cdots \alpha_{q-1}\beta\alpha_{q+1} \cdots \alpha_n) = \pm\tfrac{1}{2}n(n-1)c(\alpha_1 \cdots \alpha_n),$$

$$\sum_{\alpha\beta} \sum_{p<q}^{n} (\alpha_p\alpha_q\, |I_{\text{elec}}|\, \alpha\beta) \times c(\alpha_1 \cdots \alpha_{p-1}\alpha\alpha_{p+1} \cdots \alpha_{q-1}\beta\alpha_{q+1} \cdots \alpha_n) = -\tfrac{1}{2}n(n-1)c(\alpha_1 \cdots \alpha_n). \quad (7)$$

The space of all n-atom wavefunctions $c(\alpha_1 \cdots \alpha_n)$ satisfying these subsidiary conditions is completely equivalent to the space of all n-nucleus, ln-electron wavefunctions $\psi(X_1 \cdots X_n\, x_1 \cdots x_{ln})$ (with proper statistics).

The Schrödinger equation for ψ is

$$H\psi \equiv \left[\sum_{j=1}^{n} T(X_j) + \sum_{j=1}^{ln} T(x_j) + \sum_{j<k}^{n} V(X_jX_k) + \sum_{j<k}^{ln} V(x_jx_k) + \sum_{j=1}^{n}\sum_{k=1}^{ln} V(X_jx_k)\right]\psi = i\hbar\,\partial\psi/\partial t, \quad (8)$$

where the single-particle operators T represent the nuclear and electron kinetic energies plus any external fields, and the two-particle operators V are the nucleus–nucleus, electron–electron, and nucleus–electron Coulomb interactions. This transforms into the representation in terms of the many-atom wavefunctions[5] $c(\alpha_1 \cdots \alpha_n)$ as

$$(H - i\hbar\,\partial/\partial t)c(\alpha_1 \cdots \alpha_n) = 0, \qquad H = T + V_0 + V', \quad (9)$$

where T, V_0, and V' are, respectively, the kinetic

[5] The dependence of $c(\alpha_1 \cdots \alpha_n)$ on the time t is not indicated explicitly, but is to be understood.

energy, internal potential energy, and interatomic interaction energy operators defined by

$$Tc(\alpha_1 \cdots \alpha_n) = \sum_{\alpha} \sum_{p=1}^{n} (\alpha_p \,|T|\, \alpha) \times c(\alpha_1 \cdots \alpha_{p-1}\alpha\alpha_{p+1} \cdots \alpha_n),$$

$$V_0c(\alpha_1 \cdots \alpha_n) = \sum_{\alpha} \sum_{p=1}^{n} (\alpha_p \,|V|\, \alpha) \times c(\alpha_1 \cdots \alpha_{p-1}\alpha\alpha_{p+1} \cdots \alpha_n),$$

$$V'c(\alpha_1 \cdots \alpha_n) = \sum_{\alpha\beta} \sum_{p<q}^{n} (\alpha_p\alpha_q \,|V|\, \alpha\beta) \times c(\alpha_1 \cdots \alpha_{p-1}\alpha\alpha_{p+1} \cdots \alpha_{q-1}\beta\alpha_{q+1} \cdots \alpha_n). \quad (10)$$

The single-atom and interatomic-interaction matrix elements are defined by

$$(\alpha_p \,|T|\, \alpha) = \int \varphi^*_{\alpha_p}(Xx_1 \cdots x_l)[T(X) + lT(x_1)] \times \varphi_\alpha(Xx_1 \cdots x_l)\, dX\, dx_1 \cdots dx_l,$$

$$(\alpha_p \,|V|\, \alpha) = \int \varphi^*_{\alpha_p}(Xx_1 \cdots x_l)[\tfrac{1}{2}l(l-1)V(x_1x_2) + lV(Xx_1)]\varphi_\alpha(Xx_1 \cdots x_l)\, dX\, dx_1 \cdots dx_l,$$

$$(\alpha_p\alpha_q \,|V|\, \alpha\beta) = \int \varphi^*_{\alpha_p}(Xx_1 \cdots x_l)\varphi^*_{\alpha_q}(X'x'_1 \cdots x'_l) \times \{V(XX') + l^2V(x_1x'_1) + l[V(Xx'_1) + V(X'x_1)]\} \times \varphi_\alpha(Xx_1 \cdots x_l)\varphi_\beta(X'x'_1 \cdots x'_l) \times dX\, dx_1 \cdots dx_l\, dX'\, dx'_1 \cdots dx'_l. \quad (11)$$

Equations (9)–(11) are derived by expanding the various quantities $T\psi$ and $V\psi$ in terms of the φ_α, making use of the antisymmetry of these functions in the electron coordinates, and equating to zero the coefficients of like terms $\varphi_{\alpha_1}\cdots\varphi_{\alpha_n}$ in $(H - i\hbar\,\partial/\partial t)\psi$. The expansions of the terms $V\psi$ have different forms depending on whether the two arguments of V are in the same or different single-atom wavefunctions φ_α, giving rise to the terms V_0 and V', respectively, in H.

A quantized-field representation can now be introduced by any of the usual methods used for systems of *elementary* particles. The Fock representation[6] is most convenient for our purposes since it allows a unified treatment of the cases of discrete and continuous single-atom quantum numbers α. Thus we introduce state vectors of the form

$$\mathbf{c} = \begin{pmatrix} c_0 \\ c_1(\alpha_1) \\ \vdots \\ c_n(\alpha_1 \cdots \alpha_n) \\ \vdots \end{pmatrix}, \quad (12)$$

where c_0 is the vacuum amplitude, c_1 the one-atom amplitude, etc; the function $c(\alpha_1 \cdots \alpha_n)$ in (9) and (10) is then to be interpreted as a special case in which $\mathbf{c}$ is a total-atom-number eigenstate with eigenvalue n, i.e., all rows of (12) are zero except for the nth, which is $c(\alpha_1 \cdots \alpha_n)$. The inner product of two states of the form (12) is

$$(\mathbf{c}, \mathbf{c}') = c_0^*c'_0 + \sum_{n=1}^{\infty} \sum_{\alpha_1 \cdots \alpha_n} c_n^*(\alpha_1 \cdots \alpha_n)c'_n(\alpha_1 \cdots \alpha_n). \quad (13)$$

The atomic annihilation and creation operators a_α and $a_\alpha^\dagger$ are defined by

$$a_\alpha \begin{pmatrix} c_0 \\ c_1(\alpha_1) \\ \vdots \\ c_n(\alpha_1 \cdots \alpha_n) \\ \vdots \end{pmatrix} = \begin{pmatrix} c_1(\alpha) \\ 2^{\frac{1}{2}}c_2(\alpha_1\alpha) \\ \vdots \\ (n+1)^{\frac{1}{2}}c_{n+1}(\alpha_1 \cdots \alpha_n\alpha) \\ \vdots \end{pmatrix}, \quad (14)$$

and

$$a_\alpha^\dagger \begin{pmatrix} c_0 \\ c_1(\alpha_1) \\ \vdots \\ c_n(\alpha_1 \cdots \alpha_n) \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ \delta_{\alpha\alpha_1}c_0 \\ \vdots \\ (n^{\frac{1}{2}}/n!) \sum_P (\pm 1)^{p(P)} P[\delta_{\alpha\alpha_n}c_{n-1}(\alpha_1 \cdots \alpha_{n-1})] \\ \vdots \end{pmatrix}, \quad (15)$$

where $\sum_P$ runs over all permutations P of the arguments $\alpha_1 \cdots \alpha_n$, $p(P)$ being the parity of the permutation; the upper sign is to be taken for Bose statistics [$2J + l$ even; see discussion following Eq. (3)], and the lower for Fermi statistics. These annihilation and creation operators satisfy "ele-

[6] V. Fock, Z. Physik 75, 622 (1932).

mentary-particle" commutation or anticommutation relations

$$[a_\alpha, a_\beta]_- = [a_\alpha^\dagger, a_\beta^\dagger]_- = 0, \quad [a_\alpha, a_\beta^\dagger]_- = \delta_{\alpha\beta}. \tag{16}$$

Equations (12)–(16), as well as all subsequent equations, are valid for either discrete or continuous indices; in the continuum case $\sum_\alpha$ is to be interpreted as $\int d\alpha$, $\delta_{\alpha\beta}$ as a Dirac delta function $\delta(\alpha - \beta)$, etc.

A Hamiltonian of the form

$$H = \sum_{p=1}^{n} H(\alpha_p) + \sum_{p<q}^{n} H(\alpha_p \alpha_q) \tag{17}$$

acting on wavefunctions of the form $c(\alpha_1 \cdots \alpha_n)$ is transformed to the quantized-field representation by the familiar formula

$$H = \sum_\alpha a_\alpha^\dagger H(\alpha) a_\alpha + \tfrac{1}{2} \sum_{\alpha\beta} a_\alpha^\dagger a_\beta^\dagger H(\alpha\beta) a_\beta a_\alpha. \tag{18}$$

Comparing (17) with (9) and (10), one sees that, for our many-atom Hamiltonian, the single-atom and two-atom operators $H(\alpha)$ and $H(\alpha\beta)$ are integral operators whose action on arbitrary functions f_α and $f_{\alpha\beta}$ of α and β are

$$\begin{aligned} H(\alpha) f_\alpha &= \sum_\beta [(\alpha \,|T|\, \beta) + (\alpha \,|V|\, \beta)] f_\beta, \\ H(\alpha\beta) f_{\alpha\beta} &= \sum_{\gamma\delta} (\alpha\beta \,|V|\, \gamma\delta) f_{\gamma\delta}. \end{aligned} \tag{19}$$

Hence (18) becomes

$$\begin{aligned} &H = T + V_0 + V', \quad T = \sum_{\alpha\beta} (\alpha \,|T|\, \beta) a_\alpha^\dagger a_\beta, \\ &V_0 = \sum_{\alpha\beta} (\alpha \,|V|\, \beta) a_\alpha^\dagger a_\beta, \\ &V' = \tfrac{1}{2} \sum_{\alpha\beta\gamma\delta} (\alpha\beta \,|V|\, \gamma\delta) a_\alpha^\dagger a_\beta^\dagger a_\delta a_\gamma, \end{aligned} \tag{20}$$

where the matrix elements are defined by (11). It is clear on physical grounds that

$$H_0 \equiv T + V_0 \tag{21}$$

must describe a system of noninteracting atoms. In fact,

$$(\alpha \,|T|\, \beta) + (\alpha \,|V|\, \beta) = E_\alpha \delta_{\alpha\beta} \tag{22}$$

is just the necessary and sufficient condition that the φ_α be single-atom energy eigenfunctions with eigenvalues E_α; hence

$$H_0 = \sum_\alpha E_\alpha a_\alpha^\dagger a_\alpha. \tag{23}$$

H_0 is the natural starting point for a perturbation treatment of a system of atoms.

It is clear from a comparison of (7) with (10) and (20) that the nucleus and electron exchange operators I_{nuc} and I_{elec} are given in the quantized-field representation by

$$\begin{aligned} I_{\text{nuc}} &= \tfrac{1}{2} \sum_{\alpha\beta\gamma\delta} (\alpha\beta \,|I_{\text{nuc}}|\, \gamma\delta) a_\alpha^\dagger a_\beta^\dagger a_\delta a_\gamma, \\ I_{\text{elec}} &= \tfrac{1}{2} \sum_{\alpha\beta\gamma\delta} (\alpha\beta \,|I_{\text{elec}}|\, \gamma\delta) a_\alpha^\dagger a_\beta^\dagger a_\delta a_\gamma, \end{aligned} \tag{24}$$

with matrix elements given by the exchange integrals (6); the subsidiary conditions (7) selecting the physical n-atom state vectors Φ are

$$\begin{aligned} I_{\text{nuc}}\Phi &= (-1)^{2J} \tfrac{1}{2} n(n-1)\Phi, \\ I_{\text{elec}}\Phi &= -\tfrac{1}{2} n(n-1)\Phi, \end{aligned} \tag{25}$$

where J is the nuclear spin. It is readily verified with the aid of (6) that I_{nuc} and I_{elec} are Hermitian, and thus may be regarded as observables whose eigenvalues for physical n-atom states are given by (25). It can furthermore be shown that they commute with all observables,[7] and therefore establish a *superselection rule*[8]: There are no nonvanishing matrix elements of any observable connecting the subspace of physical n-atom states with unphysical subspaces.

The zero-temperature n-atom problem is that of finding the simultaneous eigenstates of H, of the total atom-number operator

$$N = \sum_\alpha a_\alpha^\dagger a_\alpha, \tag{26}$$

with eigenvalue n, and of I_{nuc} and I_{elec} with eigenvalues $(-1)^{2J} \tfrac{1}{2} n(n-1)$ and $-\tfrac{1}{2} n(n-1)$, respectively. In order to see how I_{nuc} and I_{elec} could be handled in a practical problem, and to clarify their relation to the intuitive picture of overlap forces, it is best to generalize at once to the nonzero-temperature case, deriving the thermal-equilib-

[7] It follows trivially from (24) that I_{nuc} and I_{elec} commute with the total atom-number operator N[Eq. (26)]. To prove that it commutes with other observables, it is best to return to the representation in terms of many-nucleus, many-electron wave functions $\psi(X_1 \cdots X_n x_1 \cdots x_{ln})$. According to (4), (5), and (7), I_{nuc} and I_{elec} are given in this representation by

$$\begin{aligned} &I_{\text{nuc}}\psi(X_1 \cdots X_n x_1 \cdots x_{ln}) \\ &\quad = \sum_{p<q}^{n} \psi(X_1 \cdots X_{p-1} X_q X_{p+1} \\ &\qquad \cdots X_{q-1} X_p X_{q+1} \cdots X_n x_1 \cdots x_{ln}), \\ &I_{\text{elec}}\psi(X_1 \cdots X_n x_1 \cdots x_{ln}) \\ &\quad = \sum_{p<q}^{n} \psi(X_1 \cdots X_n x_1 \cdots x_{lp-l} x_{lq-l+1} x_{lp-l+2} \\ &\qquad \cdots x_{lq-l} x_{lp-l+1} x_{lq-l+2} \cdots x_{ln}). \end{aligned}$$

Since I_{nuc} and I_{elec} operate on different arguments, they commute with each other. It is furthermore clear that they commute with any operator which is a symmetric function of the X_i and of the x_i. Since this is the case for any observable, e.g., for the Hamiltonian (8), I_{nuc} and I_{elec} commute with all observables.

[8] G. C. Wick, A. S. Wightman, and E. P. Wigner, Phys. Rev. 88, 101 (1952).

rium properties of the system from a suitable generalization of the usual grand partition function. Since the controllable[9] constants of the motion[10] are H, N, I_{nuc}, and I_{elec}, the appropriate density operator is

$$\rho = \Xi^{-1} \exp\left[-\beta(H - \mu N + \gamma_{\text{nuc}} I_{\text{nuc}} + \gamma_{\text{elec}} I_{\text{elec}})\right], \tag{27}$$

where Ξ is the generalized grand partition function:

$$\Xi = \text{Tr} \exp\left[-\beta(H - \mu N + \gamma_{\text{nuc}} I_{\text{nuc}} + \gamma_{\text{elec}} I_{\text{elec}})\right]. \tag{28}$$

Here $\beta = (\kappa T)^{-1}$ with κ Boltzmann's constant and T the absolute temperature, μ is the *atomic* chemical potential, γ_{nuc} and γ_{elec} are Lagrange multipliers associated with the subsidiary conditions (25), and the trace is over the whole space of *both physical and unphysical* many-atom states generated from the vacuum by the atomic creation operators $a_\alpha^\dagger$. The thermal average of any observable O is given by

$$\langle O \rangle = \text{Tr}\,(O\rho). \tag{29}$$

The parameters μ, γ_{nuc}, and γ_{elec} are determined by the coupled equations

$$\begin{aligned} -(\partial W/\partial \mu)_{\beta,\gamma_{\text{nuc}},\gamma_{\text{elec}}} &= \langle N \rangle = n, \\ (\partial W/\partial \gamma_{\text{nuc}})_{\beta,\mu,\gamma_{\text{elec}}} &= \langle I_{\text{nuc}} \rangle = (-1)^{2J} \tfrac{1}{2} n(n-1), \\ (\partial W/\partial \gamma_{\text{elec}})_{\beta,\mu,\gamma_{\text{nuc}}} &= \langle I_{\text{elec}} \rangle = -\tfrac{1}{2} n(n-1), \end{aligned} \tag{30}$$

where the subscripts denote, as usual, the quantities held constant in the differentiations, and W is the generalized thermodynamic potential related to Ξ by

$$\Xi = e^{-\beta W}. \tag{31}$$

The density operator ρ describes an open system in which N, I_{nuc}, and I_{elec} undergo small fluctuations about their mean values specified by the right sides of Eqs. (30). The system is, however, "open" in a more general sense than the usual one, since the fluctuations of I_{nuc} and I_{elec}, although partly due merely to the fluctuations of N, are also due partly to temporary excursions into the unphysical part of the many-atom state space. In analogy with the usual relationship between the canonical and grand canonical ensembles, one expects that the probability distributions of the eigenvalues of N, I_{nuc}, and I_{elec} will be very sharply peaked about their mean values for large n, and in fact that the properties of the system derived from (28) will agree in the limit $[n \to \infty, \Omega \to \infty, (n/\Omega) \to \text{const}]$ with those that would be derived from a generalized *canonical* partition function tr $e^{-\beta H}$ in which the trace is restricted to eigenstates of N with eigenvalue n which are also *physical* n-atom states, i.e., satisfy (25) as *eigenvalue* equations. In fact, it is shown in the Appendix that the eigenvalues of I_{nuc} and I_{elec} for n-atom states all lie on the interval $[-\frac{1}{2}n\,(n-1), +\frac{1}{2}n\,(n-1)]$; hence if the fluctuations of N are negligible and the mean-value conditions (30) are satisfied to leading order in n, *then the fluctuations of each I will also be negligible*, since this is the only way the *mean* value could lie on one *end* of the range of eigenvalues.

The terms $\gamma_{\text{nuc}} I_{\text{nuc}}$ and $\gamma_{\text{elec}} I_{\text{elec}}$ in (28) have the same general structure as the interatomic interaction V' [compare Eqs. (20) and (24)]. This is the mathematical expression, for the many-atom problem, of the familiar picture of the strong short-range interatomic repulsion as arising from overlap of electronic shells. One expects on physical grounds that, at low and moderate temperatures and pressures, the nuclear exchange interaction will be negligible compared to the electron exchange interaction so that $|\gamma_{\text{nuc}}| \ll |\gamma_{\text{elec}}|$, and furthermore that the electron exchange interaction, and hence γ_{elec}, will be almost independent of temperature. At temperatures and pressures high enough that interpenetration of atoms is probable, one expects that γ_{nuc} and γ_{elec} will be of the same order of magnitude and both will be temperature-dependent.

We conclude this section by examining the nature of the coordinate-space interatomic interactions implied by the structures of the Coulomb interatomic interaction Hamiltonian V' and the exchange interatomic interaction "Hamiltonians" $\gamma_{\text{nuc}} I_{\text{nuc}}$ and $\gamma_{\text{elec}} I_{\text{elec}}$. This is facilitated by introduction of annihilation and creation operators for localized atoms by a generalization of the usual procedure for elementary bosons or fermions. We denote by $(\mathbf{k}\bar{\alpha})$ that particular single-atom index α which consists of the center-of-mass translational wave vector $\mathbf{k}$ and the set $\bar{\alpha}$ of all internal quantum numbers necessary to specify the atomic wavefunction in the center-of-mass system, and define

$$\psi_{\bar{\alpha}}(\mathbf{r}) \equiv \Omega^{-\frac{1}{2}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} a_{\mathbf{k}\bar{\alpha}} \tag{32}$$

where Ω is the volume of the system; the allowed values of $\mathbf{k}$ are determined by periodic boundary conditions (all components of $\mathbf{k}$ are integral multiples of $2\pi\Omega^{-\frac{1}{3}}$). Then it follows from (16) that

[9] A. I. Khinchin, *Mathematical Foundations of Statistical Mechanics* (Dover Publications, Inc., New York, 1949), p. 51.

[10] In some cases one would want to introduce additional commuting and controllable observables, e.g., the total linear momentum. The necessary modifications in such cases are obvious.

$$[\psi_{\bar{\alpha}}(\mathbf{r}), \psi_{\bar{\beta}}(\mathbf{r}')]_{\mp} = [\psi^{\dagger}_{\bar{\alpha}}(\mathbf{r}), \psi^{\dagger}_{\bar{\beta}}(\mathbf{r}')]_{\mp} = 0,$$

$$[\psi_{\bar{\alpha}}(\mathbf{r}), \psi^{\dagger}_{\bar{\beta}}(\mathbf{r}')]_{\mp} = \Omega^{-1} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \delta_{\bar{\alpha}\bar{\beta}} = \delta(\mathbf{r}-\mathbf{r}')\delta_{\bar{\alpha}\bar{\beta}}. \quad (33)$$

These differ from the commutation relations for elementary bosons or fermions only in that $\bar{\alpha}$ and $\bar{\beta}$ contain other quantum numbers in addition to spins; thus the atoms behave in this representation like *point* particles with internal degrees of freedom. The inverse transformation to (32) is

$$a_{\mathbf{k}\bar{\alpha}} = \Omega^{-\frac{1}{2}} \int d^3\mathbf{r} e^{-i\mathbf{k}\cdot\mathbf{r}} \psi_{\bar{\alpha}}(\mathbf{r}). \quad (34)$$

Substitution into (20) gives

$$V' = \tfrac{1}{2} \sum_{\bar{\alpha}\bar{\beta}\bar{\gamma}\bar{\delta}} \int d^3\mathbf{r}_1 \cdots d^3\mathbf{r}_4 v_{\bar{\alpha}\bar{\beta}\bar{\gamma}\bar{\delta}}(\mathbf{r}_1 \cdots \mathbf{r}_4) \times \psi^{\dagger}_{\bar{\alpha}}(\mathbf{r}_1)\psi^{\dagger}_{\bar{\beta}}(\mathbf{r}_2)\psi_{\bar{\delta}}(\mathbf{r}_4)\psi_{\bar{\gamma}}(\mathbf{r}_3), \quad (35)$$

with

$$v_{\bar{\alpha}\bar{\beta}\bar{\gamma}\bar{\delta}}(\mathbf{r}_1 \cdots \mathbf{r}_4) = \Omega^{-2} \sum_{\mathbf{k}_1 \cdots \mathbf{k}_4} (\mathbf{k}_1\bar{\alpha}, \mathbf{k}_2\bar{\beta} \,|V|\, \mathbf{k}_3\bar{\gamma}, \mathbf{k}_4\bar{\delta}) \times e^{i(\mathbf{k}_1\cdot\mathbf{r}_1 + \mathbf{k}_2\cdot\mathbf{r}_2 - \mathbf{k}_3\cdot\mathbf{r}_3 - \mathbf{k}_4\cdot\mathbf{r}_4)}, \quad (36)$$

Similarly, it follows from (24) that the exchange interactions $\gamma_{\text{nuc}} I_{\text{nuc}}$ and $\gamma_{\text{elec}} I_{\text{elec}}$ are given by expressions differing from (35) only through replacement of $v_{\bar{\alpha}\bar{\beta}\bar{\gamma}\bar{\delta}}$ by $v^{\text{nuc}}_{\bar{\alpha}\bar{\beta}\bar{\gamma}\bar{\delta}}$ or $v^{\text{elec}}_{\bar{\alpha}\bar{\beta}\bar{\gamma}\bar{\delta}}$, defined by replacing V in (36) by $\gamma_{\text{nuc}} I_{\text{nuc}}$ or $\gamma_{\text{elec}} I_{\text{elec}}$.

In the case of real atoms, the potentials V in (11) are Coulomb interactions;

$$\begin{aligned} V(X_iX_k) &= (Ze)^2/|\mathbf{R}_i - \mathbf{R}_k|, \\ V(x_ix_k) &= e^2/|\mathbf{r}_i - \mathbf{r}_k|, \\ V(X_ix_k) &= -Ze^2/|\mathbf{R}_i - \mathbf{r}_k|. \end{aligned} \quad (37)$$

To simplify (35), we introduce the center-of-mass decomposition of the atomic wavefunctions;

$$\varphi_{\mathbf{k}\bar{\alpha}}(Xx_1 \cdots x_l) = \Omega^{-\frac{1}{2}} e^{i\mathbf{k}\cdot\mathbf{R}} \tilde{\varphi}_{\bar{\alpha}}(Xx_1 \cdots x_l), \quad (38)$$

where $\tilde{\varphi}_{\bar{\alpha}}$ is the wavefunction in the center-of-mass system; to avoid nonessential complications we neglect finite-nuclear-mass corrections and hence replace the center-of-mass position $\mathbf{R}_{\text{c.m.}}$ by the nuclear position $\mathbf{R}$ in the translational wavefunction $\Omega^{-\frac{1}{2}} e^{i\mathbf{k}\cdot\mathbf{R}_{\text{c.m.}}}$. A straightforward derivation making use of translational invariance of the $\tilde{\varphi}_{\bar{\alpha}}$, completeness of the exponentials, and the relation $l = Z$ following from electrical neutrality gives

$$V' = \tfrac{1}{2} \sum_{\bar{\alpha}\bar{\beta}\bar{\gamma}\bar{\delta}} \int d^3\mathbf{r} d^3\mathbf{r}' \psi^{\dagger}_{\bar{\alpha}}(\mathbf{r})\psi^{\dagger}_{\bar{\beta}}(\mathbf{r}') \times v_{\bar{\alpha}\bar{\beta}\bar{\gamma}\bar{\delta}}(\mathbf{r}-\mathbf{r}')\psi_{\bar{\delta}}(\mathbf{r}')\psi_{\bar{\gamma}}(\mathbf{r}), \quad (39)$$

with

$$\begin{aligned} v_{\bar{\alpha}\bar{\beta}\bar{\gamma}\bar{\delta}}(\mathbf{r}-\mathbf{r}') &= (Ze)^2 \int \tilde{\varphi}^*_{\bar{\alpha}}(Xx_1 \cdots x_l)\tilde{\varphi}^*_{\bar{\beta}}(X'x'_1 \cdots x'_l) \\ &\times (|\mathbf{R} - \mathbf{R}' + \mathbf{r} - \mathbf{r}'|^{-1} + |\mathbf{r} - \mathbf{r}'_1 + \mathbf{r} - \mathbf{r}'|^{-1} \\ &- |\mathbf{R} - \mathbf{r}'_1 + \mathbf{r} - \mathbf{r}'|^{-1} - |\mathbf{r}_1 - \mathbf{R}' + \mathbf{r} - \mathbf{r}'|^{-1}) \\ &\times \tilde{\varphi}_{\bar{\gamma}}(Xx_1 \cdots x_l)\tilde{\varphi}_{\bar{\delta}}(X'x'_1 \cdots x'_l)\delta(\mathbf{R})\delta(\mathbf{R}') \\ &\times dX\, dx_1 \cdots dx_l\, dX'\, dx'_1 \cdots dx'_l. \end{aligned} \quad (40)$$

Equation (39) has the same structure as the usual expression for the two-body interaction Hamiltonian in position space, except that because of the internal degrees of freedom of the atoms the usual interaction potential $v(\mathbf{r} - \mathbf{r}')$ is replaced here by a potential matrix $v_{\bar{\alpha}\bar{\beta}\bar{\gamma}\bar{\delta}}(\mathbf{r} - \mathbf{r}')$, indexed by the initial and final internal quantum numbers of the pair of interacting atoms; the elements diagonal in these indices (those with $\bar{\gamma} = \bar{\alpha}$ and $\bar{\delta} = \bar{\beta}$) describe elastic scattering whereas the off-diagonal elements describe inelastic scattering. The exact expression for V' including finite-nuclear-mass corrections differs only in that $\delta(\mathbf{R})$ and $\delta(\mathbf{R}')$ in (40) are replaced by $\delta(\mathbf{R}_{\text{c.m.}})$ and $\delta(\mathbf{R}'_{\text{c.m.}})$. A similar derivation gives for the nuclear exchange interaction

$$\gamma_{\text{nuc}} I_{\text{nuc}} = \tfrac{1}{2} \sum_{\bar{\alpha}\bar{\beta}\bar{\gamma}\bar{\delta}} \int d^3\mathbf{r}\, d^3\mathbf{r}' \psi^{\dagger}_{\bar{\alpha}}(\mathbf{r})\psi^{\dagger}_{\bar{\beta}}(\mathbf{r}') \times v^{\text{nuc}}_{\bar{\alpha}\bar{\beta}\bar{\gamma}\bar{\delta}}(\mathbf{r}-\mathbf{r}')\psi_{\bar{\delta}}(\mathbf{r})\psi_{\bar{\gamma}}(\mathbf{r}'), \quad (41)$$

with

$$\begin{aligned} v^{\text{nuc}}_{\bar{\alpha}\bar{\beta}\bar{\gamma}\bar{\delta}}(\mathbf{r}-\mathbf{r}') &= \gamma_{\text{nuc}} \int \tilde{\varphi}^*_{\bar{\alpha}}(Xx_1 \cdots x_l) \\ &\times \tilde{\varphi}^*_{\bar{\beta}}(X'x'_1 \cdots x'_l)\tilde{\varphi}_{\bar{\gamma}}(X' + \mathbf{r}' - \mathbf{r}, x_1 \cdots x_l) \\ &\times \tilde{\varphi}_{\bar{\delta}}(X + \mathbf{r} - \mathbf{r}', x'_1 \cdots x'_l)\delta(\mathbf{R})\delta(\mathbf{R}') \\ &\times dX\, dx_1 \cdots dx_l\, dX'\, dx'_1 \cdots dx'_l, \end{aligned} \quad (42)$$

where, e.g., $X + \mathbf{r} - \mathbf{r}'$ is a symbolic expression for $(\mathbf{R} + \mathbf{r} - \mathbf{r}', \sigma)$ where $X = (\mathbf{R}, \sigma)$, σ being the nuclear spin variable which is summed over as part of $\int dX$. The electron-exchange interaction is

$$\gamma_{\text{elec}} I_{\text{elec}} = \tfrac{1}{2} \sum_{\bar{\alpha}\bar{\beta}\bar{\gamma}\bar{\delta}} \int d^3\mathbf{r}\, d^3\mathbf{r}' \psi^{\dagger}_{\bar{\alpha}}(\mathbf{r})\psi^{\dagger}_{\bar{\beta}}(\mathbf{r}') \times v^{\text{elec}}_{\bar{\alpha}\bar{\beta}\bar{\gamma}\bar{\delta}}(\mathbf{r}-\mathbf{r}')\psi_{\bar{\delta}}(\mathbf{r}')\psi_{\bar{\gamma}}(\mathbf{r}), \quad (43)$$

with

$$\begin{aligned} v^{\text{elec}}_{\bar{\alpha}\bar{\beta}\bar{\gamma}\bar{\delta}}(\mathbf{r}-\mathbf{r}') &= \gamma_{\text{elec}} \int \tilde{\varphi}^*_{\bar{\alpha}}(Xx_1 \cdots x_l) \\ &\times \tilde{\varphi}^*_{\bar{\beta}}(X'x'_1 \cdots x'_l)\tilde{\varphi}_{\bar{\gamma}}(X, x'_1 + \mathbf{r}' - \mathbf{r}, x_2 \cdots x_l) \\ &\times \tilde{\varphi}_{\bar{\delta}}(X', x_1 + \mathbf{r} - \mathbf{r}', x'_2 \cdots x'_l)\delta(\mathbf{R})\delta(\mathbf{R}') \\ &\times dX\, dx_1 \cdots dx_l\, dX'\, dx'_1 \cdots dx'_l. \end{aligned} \quad (44)$$

It is to be noted that the arguments of $\psi_{\bar{\gamma}}$ and $\psi_{\bar{\delta}}$ in (41) are interchanged relative to their order in (39), whereas they have the normal order in (43). If finite-nuclear-mass corrections are included then the nuclear and electron exchange interactions can no longer be exactly written in the forms (41) and (43), but have the more general structure of (35); nevertheless, $v^{\text{nuc}}_{\bar{\alpha}\bar{\beta}\bar{\gamma}\bar{\delta}}(\mathbf{r}_1 \cdots \mathbf{r}_4)$ is sharply peaked about $\mathbf{r}_3 = \mathbf{r}_1$ and $\mathbf{r}_4 = \mathbf{r}_2$, and $v^{\text{elec}}_{\bar{\alpha}\bar{\beta}\bar{\gamma}\bar{\delta}}(\mathbf{r}_1 \cdots \mathbf{r}_4)$ about $\mathbf{r}_3 = \mathbf{r}_2$ and $\mathbf{r}_4 = \mathbf{r}_1$.

The potential matrix $v_{\bar{\alpha}\bar{\beta}\bar{\gamma}\bar{\delta}}(\mathbf{r} - \mathbf{r}')$ does not directly contain the contributions of dispersion forces such as the van der Waals interaction; instead, as in more familiar formulations,[11] these arise from certain terms of second and higher order in V' in an evaluation of the energy. In order to make this point clearer let us consider the leading terms in a perturbation calculation of the ground-state energy. The unperturbed n-atom ground state can be written in the Bose[12] case as

$$\Phi_0 = (n!)^{-\frac{1}{2}}(a_0^\dagger)^n\,|0), \tag{45}$$

where $|0)$ is the vacuum [corresponding to a state vector $\mathbf{c}$ in (12) for which $c_0 = 1$ and all other c_i vanish], and $a_0^\dagger$ creates an atom in the single-atom ground state $\varphi_0(Xx_1 \cdots x_l)$. It then follows from (39), (32), and (16) that the first-order perturbation energy is just

$$(\Phi_0,\ V'\Phi_0) = \tfrac{1}{2}n(n-1)\Omega^{-1}\int v_{0000}(\mathbf{r})\,d^3r, \tag{46}$$

the number of atomic pairs times the space-averaged static Coulomb interaction energy of one pair; this is a repulsive energy and does not contain any of the van der Waals interaction. In second order there are two types of excitation processes to be considered: the elastic processes in which two atoms in their ground states φ_0 interact and emerge with nonzero equal and opposite momenta $\pm\mathbf{k}$ but still in their *internal* ground states, and the inelastic processes in which the atoms emerge not only with momenta $\pm\mathbf{k}$ but also *internally* excited. The elastic processes give a correction to (46) which is of the same form as the second-order perturbation energy for a system of structureless particles, and need not be considered further here; the inelastic processes give the van der Waals interaction. Indeed, expanding $v_{\bar{\alpha}\bar{\beta}00}$ in inverse powers of $|\mathbf{r} - \mathbf{r}'|$, one finds

$$v_{\bar{\alpha}\bar{\beta}00}(\mathbf{r} - \mathbf{r}') = (Ze)^2\int \bar{\varphi}^*_{\bar{\alpha}}\bar{\varphi}^*_{\bar{\beta}}\left\{\frac{(\mathbf{r}_1 - \mathbf{R})\cdot(\mathbf{r}_1' - \mathbf{R}')}{|\mathbf{r} - \mathbf{r}'|^3} - 3\,\frac{[(\mathbf{r} - \mathbf{r}')\cdot(\mathbf{r}_1 - \mathbf{R})][(\mathbf{r} - \mathbf{r}')\cdot(\mathbf{r}_1' - \mathbf{R}')]}{|\mathbf{r} - \mathbf{r}'|^5}\right\} \times \bar{\varphi}_0\bar{\varphi}_0\delta\delta + \cdots, \tag{47}$$

where the integration variables and arguments of the $\bar{\varphi}$ and δ functions are the same as in (40). The factor

$$(Ze)^2\left\{\frac{(\mathbf{r}_1 - \mathbf{R})\cdot(\mathbf{r}_1' - \mathbf{R}')}{|\mathbf{r} - \mathbf{r}'|^3} - 3\,\frac{[(\mathbf{r} - \mathbf{r}')\cdot(\mathbf{r}_1 - \mathbf{R})][(\mathbf{r} - \mathbf{r}')\cdot(\mathbf{r}_1' - \mathbf{R}')]}{|\mathbf{r} - \mathbf{r}'|^5}\right\}$$

is clearly the instantaneous dipole–dipole interaction taking into account the indistinguishability of all electrons in a given atom. The expression (47) vanishes identically unless $\bar{\varphi}_{\bar{\alpha}}$ and $\bar{\varphi}_{\bar{\beta}}$ both have odd parity (we assume that $\bar{\varphi}_0$ has even parity), in which case the corresponding contribution to the second-order perturbation energy is nonzero, negative (by a general theorem on second-order perturbation contributions), and varies like $|\mathbf{r} - \mathbf{r}'|^{-6}$, the usual van der Waals interaction.

In a consistent calculation to a given order of perturbation theory, it would, of course, be necessary to consider $(V' + \gamma_{\text{nuc}}I_{\text{nuc}} + \gamma_{\text{elec}}I_{\text{elec}})$ as the perturbation, rather than merely V', in order that the last two of Eqs. (30) be satisfied (to given order) by the *perturbed* state vector. Thus the exchange interactions also lead to dispersion forces. It follows from the general properties of bound-state wavefunctions that $v^{\text{nuc}}_{\bar{\alpha}\bar{\beta}00}(\mathbf{r} - \mathbf{r}')$ and $v^{\text{elec}}_{\bar{\alpha}\bar{\beta}00}(\mathbf{r} - \mathbf{r}')$ fall off exponentially for large $|\mathbf{r} - \mathbf{r}'|$, and hence are negligible compared to the van der Waals interaction at large distances. In an actual calculation using the many-atom formalism of this paper one need not (and usually should not) make any explicit calculation of the dispersion forces, but in treating physical situations where dispersion forces are not negligible (e.g., liquid He^3 or He^4), it would be important to include those terms in $(V' + \gamma_{\text{nuc}}I_{\text{nuc}} + \gamma_{\text{elec}}I_{\text{elec}})$ which are off-diagonal in the internal indices $\bar{\alpha}$, $\bar{\beta}$, $\bar{\gamma}$, $\bar{\delta}$, or at least the matrix elements from the ground state to the first (internally) excited state.

3. ALTERNATIVE DERIVATION BY DYSON'S METHOD

In this section we shall give an alternative derivation of the atomic second-quantization formalism by a method analogous to that developed by Dyson[3] for treating the problem of spin-wave interactions,

[11] See e.g., L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1949) pp. 174 ff.

[12] The calculations are more complicated in the Fermi case, but the complications are associated with *translational* degrees of freedom, whereas the van der Waals interaction is associated with virtual excitation of *internal* degrees of freedom.

subsequently adapted to the fermion-pair "quasi-boson" problem by Blatt and Matsubara.[4] This alternative derivation is valuable both because it establishes contact with past work,[3,4] and because it furnishes added insight into the significance of the exchange interactions $\gamma_{\text{nuc}}I_{\text{nuc}}$ and $\gamma_{\text{elec}}I_{\text{elec}}$.

We begin by introducing the usual annihilation and creation operators $\Psi(X)$ and $\Psi^\dagger(X)$ for localized nuclei, and $\Psi(x)$ and $\Psi^\dagger(x)$ for localized electrons, which satisfy the commutation relations

$$[\Psi(X)\Psi(X') - (-1)^{2i}\Psi(X')\Psi(X)] = 0,$$
$$[\Psi(X)\Psi^\dagger(X') - (-1)^{2i}\Psi^\dagger(X')\Psi(X)] = \delta(X - X'),$$
$$[\Psi(x), \Psi(x')]_+ = 0, \quad [\Psi(x), \Psi^\dagger(x')]_+ = \delta(x - x'),$$
$$[\Psi(X), \Psi(x)]_- = [\Psi(X), \Psi^\dagger(x)]_- = 0, \tag{48}$$

and their Hermitian conjugates. The nucleus–electron vacuum $|0\rangle$ is defined by

$$\langle 0 \mid 0\rangle = 1, \qquad \Psi(X)\,|0\rangle = \Psi(x)\,|0\rangle = 0. \tag{49}$$

The single-atom states $\varphi_\alpha(Xx_1 \cdots x_l)$ transform into this representation as state vectors $|\alpha\rangle$ of the form

$$|\alpha\rangle = A_\alpha^\dagger\,|0\rangle, \tag{50}$$

where the *physical atom creation operator* $A_\alpha^\dagger$ is defined as

$$A_\alpha^\dagger \equiv (l!)^{-\frac{1}{2}} \int dX\, dx_1 \cdots dx_l\, \varphi_\alpha(Xx_1 \cdots x_l) \times \Psi^\dagger(X)\Psi^\dagger(x_1) \cdots \Psi^\dagger(x_l). \tag{51}$$

The normalization constant $(l!)^{-\frac{1}{2}}$ is chosen so that the states $|\alpha\rangle$ are properly orthonormal:

$$\langle \alpha \mid \beta\rangle = \langle 0|\, A_\alpha A_\beta^\dagger\,|0\rangle = \delta_{\alpha\beta}, \tag{52}$$

the annihilation operators A_α being *defined* as $A_\alpha \equiv (A_\alpha^\dagger)^\dagger$; Eq. (52) is readily checked by evaluating the vacuum expectation value by Wick's theorem[13] and using the antisymmetry of the φ_α in the electron variables.

The natural definition of an *atomic product state* representing n noninteracting atoms is now

$$A_{\alpha_1}^\dagger \cdots A_{\alpha_n}^\dagger\,|0\rangle. \tag{53}$$

It is rather obvious (and will presently be proved) that the set of all such states spans the space of all n-nucleus, ln-electron states.[14] In fact, for $n \geq 2$ the set of all n-atom product states will turn out to be redundant, i.e., *over*complete, and nonorthogonal. Although

$$[A_\alpha, A_\beta]_- = [A_\alpha^\dagger, A_\beta^\dagger]_- = 0, \tag{54}$$

where the commutator or anticommutator is to be taken depending upon whether $2J + l$ is even or odd, the annihilation operators A_α do not satisfy simple Bose or Fermi commutation relations with the creation operators $A_\beta^\dagger$; this is responsible for the nonorthogonality and overcompleteness of the atomic product states (53), and causes serious mathematical difficulties in working with these states. The situation is very similar to that encountered in the theory of spin-wave interactions.[3] There is an obvious physical definition of spin-wave creation operators, but the spin waves are not bosons and the spin-wave product states form a nonorthogonal and overcomplete set. Dyson solved the overcompleteness problem by deriving an exact formula for the partition function as a sum over a nonorthogonal and overcomplete set. We shall proceed in a different fashion, by introducing subsidiary conditions which remove the redundancy, a method which is much simpler for our problem and also directly related to the derivation in Sec. 2. Dyson solved the problem of the complicated commutation relations by introducing an *ideal state space*, in one–one correspondence with the space of physical product states, in which the commutation relations were of the simple Bose form, and in which all calculations could be performed once the formula transforming the Hamiltonian in the physical state space into that in the ideal state space was derived. In this we shall follow Dyson, introducing a space spanned by *ideal atomic product states* in which the ideal atomic annihilation and creation operators satisfy simple Bose or Fermi commutation relations. It will in fact turn out that this ideal state space is identical with that employed in Sec. 2.

We start by showing that the set of all physical n-atom product states (53) is not a linearly independent set, i.e. that there exist linear relations between these states; it will then follow from their sufficiency (which will be proved) for spanning the n-nucleus, ln-electron space that they form an *over*complete set, since if they were merely *complete* they would be linearly independent. Consider a general linear combination Φ of n-atom product states:

$$\Phi = \sum_{\alpha_1 \cdots \alpha_n} c(\alpha_1 \cdots \alpha_n) A_{\alpha_1}^\dagger \cdots A_{\alpha_n}^\dagger\,|0\rangle, \tag{55}$$

where the coefficient function c is symmetric or antisymmetric depending upon whether $2J + l$ is even

[13] A. Houriet and A. Kind, Helv. Phys. Acta **22**, 319 (1949); G. C. Wick, Phys. Rev. **80**, 268 (1950); F. J. Dyson, Phys. Rev. **82**, 428 (1951).

[14] As always, the definition of completeness is not absolute, but relative to the boundary and regularity conditions.

or odd.[15] We shall show that in general there exists another coefficient function d, with $d(\alpha_1 \cdots \alpha_n) \not\equiv c(\alpha_1 \cdots \alpha_n)$, leading to the *same* state Φ. We proceed by construction. Pick any pair (p, q) with $1 \leq p < q \leq n$, write out the explicit expression for $A^\dagger_{\alpha_p} A^\dagger_{\alpha_q}$ following from (51), and interchange the two nucleus creation operators; this gives

$$A^\dagger_{\alpha_p} A^\dagger_{\alpha_q} = (-1)^{2J}(l!)^{-1}$$
$$\times \int dX\, dx_1 \cdots dx_l\, dX'\, dx_1' \cdots dx_l'\, \varphi_{\alpha_p}(Xx_1 \cdots x_l)$$
$$\times \varphi_{\alpha_q}(X'x_1' \cdots x_l')\Psi^\dagger(X')\Psi^\dagger(x_1) \cdots \Psi^\dagger(x_l)$$
$$\times \Psi^\dagger(X)\Psi^\dagger(x_1') \cdots \Psi^\dagger(x_l'). \quad (56)$$

Next, use completeness of the φ_α (hence of the φ^*_α) to expand

$$\Psi^\dagger(X')\Psi^\dagger(x_1) \cdots \Psi^\dagger(x_l)$$
$$= \sum_\alpha \varphi^*_\alpha(X'x_1 \cdots x_l) \int \varphi_\alpha(Yy_1 \cdots y_l)$$
$$\times \Psi^\dagger(Y)\Psi^\dagger(y_1) \cdots \Psi^\dagger(y_l)\, dY\, dy_1 \cdots dy_l$$
$$= (l!)^{-\frac{1}{2}} \sum_\alpha \varphi^*_\alpha(X'x_1 \cdots x_l)A^\dagger_\alpha,$$
$$\Psi^\dagger(X)\Psi^\dagger(x_1') \cdots \Psi^\dagger(x_l')$$
$$= (l!)^{-\frac{1}{2}} \sum_\beta \varphi^*_\beta(Xx_1' \cdots x_l')A^\dagger_\beta. \quad (57)$$

Thus, by (6),

$$A^\dagger_{\alpha_p} A^\dagger_{\alpha_q} = (-1)^{2J} \sum_{\alpha\beta} (\alpha\beta\, |I_{\text{nuc}}|\, \alpha_p\alpha_q)A^\dagger_\alpha A^\dagger_\beta. \quad (58)$$

Finally, in the expression (55) for Φ, move $A^\dagger_{\alpha_q}$ to the left until it is adjacent to $A^\dagger_{\alpha_p}$, thus introducing a factor $(\pm 1)^{q-p-1}$, substitute (58), interchange the dummy indices $(\alpha\beta)$ with $(\alpha_p\alpha_q)$, and move $A^\dagger_{\alpha_q}$ back to its original position, thus canceling the factor $(\pm 1)^{q-p-1}$; this gives

$$\Phi = \sum_{\alpha_1 \cdots \alpha_n} d(\alpha_1 \cdots \alpha_n)A^\dagger_{\alpha_1} \cdots A^\dagger_{\alpha_n} |0\rangle, \quad (59)$$

with

$$d(\alpha_1 \cdots \alpha_n) = (-1)^{2J} \sum_{\alpha\beta} (\alpha_p\alpha_q\, |I_{\text{nuc}}|\, \alpha\beta)$$
$$\times c(\alpha_1 \cdots \alpha_{p-1}\alpha\alpha_{p+1} \cdots \alpha_{q-1}\beta\alpha_{q+1} \cdots \alpha_n). \quad (60)$$

Since $d(\alpha_1 \cdots \alpha_n) \not\equiv c(\alpha_1 \cdots \alpha_n)$ *in general*, this shows immediately that the set of all n-atom product states (53) is not linearly independent,[16] as is in fact already clear from (58). One can similarly show that

$$\Phi = \sum_{\alpha_1 \cdots \alpha_n} e(\alpha_1 \cdots \alpha_n)A^\dagger_{\alpha_1} \cdots A^\dagger_{\alpha_n} |0\rangle \quad (61)$$

with

$$e(\alpha_1 \cdots \alpha_n) = -\sum_{\alpha\beta} (\alpha_p\alpha_q\, |I_{\text{elec}}|\, \alpha\beta)$$
$$\times c(\alpha_1 \cdots \alpha_{p-1}\alpha\alpha_{p+1} \cdots \alpha_{q-1}\beta\alpha_{q+1} \cdots \alpha_n), \quad (62)$$

and *in general* $e(\alpha_1 \cdots \alpha_n) \not\equiv c(\alpha_1 \cdots \alpha_n)$. Other similar relationships, found by interchanging more than one pair of creation operators between $A^\dagger_{\alpha_p}$ and $A^\dagger_{\alpha_q}$ before expanding as in (57), or even by permuting creation operators between more than two $A^\dagger_\alpha$ operators, are not independent, since any permutation is a product of interchanges.

For the special case of functions $c(\alpha_1 \cdots \alpha_n)$ satisfying the subsidiary conditions (5) [or equivalently (7)], the above construction of linear relationships between the atomic product states fails because the functions $d(\alpha_1 \cdots \alpha_n)$ and $e(\alpha_1 \cdots \alpha_n)$ turn out to be identically equal to $c(\alpha_1 \cdots \alpha_n)$. One may therefore anticipate that these subsidiary conditions may just suffice to remove the redundancy of the expansion coefficients in an expansion such as (55), but without destroying the completeness of the set of all such states for which the expansion coefficients satisfy the subsidiary conditions. We now prove this conjecture. We first prove completeness, i.e. that any n-nucleus, ln-electron state Φ has an expansion of the form (55) with a coefficient function $c(\alpha_1 \cdots \alpha_n)$ satisfying the subsidiary conditions (7). One knows that any n-nucleus, ln-electron state can be expanded in the form

$$\Phi = \int dX_1 \cdots dX_n\, dx_1 \cdots dx_{ln}\, \psi(X_1 \cdots X_n x_1 \cdots x_{ln})$$
$$\times \Psi^\dagger(X_1) \cdots \Psi^\dagger(X_n)\Psi^\dagger(x_1) \cdots \Psi^\dagger(x_{ln}) |0\rangle, \quad (63)$$

where ψ is the Schrödinger wave function. But if $\{\varphi_\alpha\}$ is any complete set of one-atom wavefunctions satisfying the same boundary and regularity conditions as ψ, then ψ can be expanded in terms of products of the φ_α according to (2), with coefficients $c(\alpha_1 \cdots \alpha_n)$ given by (3). Substituting (2) into (63), making an appropriate permutation of the $\psi^\dagger$ operators, and using (51), one obtains precisely (55) apart from an additional factor $(l!)^{\frac{1}{2}n}$ and a possible overall minus sign; these factors can be absorbed by renormalization of $c(\alpha_1 \cdots \alpha_n)$. It is a trivial exercise to verify that the resultant $c(\alpha_1 \cdots \alpha_n)$ satisfies the subsidiary conditions (7), which were in fact originally derived from (4) and (2). Finally,

[15] This symmetry or antisymmetry assumption is not *essential* since, because of (54), only the symmetric or antisymmetric part of c survives the summation. However, we shall assume symmetry or antisymmetry because this greatly simplifies the algebra.

[16] One has the dependence relation

$$\sum_{\alpha_1 \cdots \alpha_n} [c(\alpha_1 \cdots \alpha_n) - d(\alpha_1 \cdots \alpha_n)]$$
$$\times A^\dagger_{\alpha_1} \cdots A^\dagger_{\alpha_n}|0\rangle = \Phi - \Phi = 0$$

with $c(\alpha_1 \cdots \alpha_n) - d(\alpha_1 \cdots \alpha_n) \not\equiv 0$.

we prove uniqueness, i.e. that for given Φ there exists only one coefficient function $c(\alpha_1 \cdots \alpha_n)$ satisfying both (7) and (55). We proceed in the standard fashion for uniqueness proofs, except for the complications entailed by the lack of independence of the states (53). Thus we suppose (7) and (55) to be satisfied by both $c(\alpha_1 \cdots \alpha_n)$ and $c'(\alpha_1 \cdots \alpha_n)$, and let $\Delta(\alpha_1 \cdots \alpha_n)$ be the difference between c and c'. Then Δ satisfies (7), and furthermore,

$$\sum_{\alpha_1 \cdots \alpha_n} \Delta(\alpha_1 \cdots \alpha_n) A^\dagger_{\alpha_1} \cdots A^\dagger_{\alpha_n} |0\rangle = 0; \quad (64)$$

we shall show that then $\Delta(\alpha_1 \cdots \alpha_n) \equiv 0$, i.e. $c(\alpha_1 \cdots \alpha_n) \equiv c'(\alpha_1 \cdots \alpha_n)$. Insert (51) into (64); this gives

$$\int dX_1 \cdots dX_n\, dx_1 \cdots dx_{ln} \Big[\sum_{\alpha_1 \cdots \alpha_n} \Delta(\alpha_1 \cdots \alpha_n) \times \varphi_{\alpha_1}(X_1 x_1 \cdots x_l) \cdots \varphi_{\alpha_n}(X_n x_{ln-l+1} \cdots x_{ln})\Big] \times \Psi^\dagger(X_1) \cdots \Psi^\dagger(X_n)\Psi^\dagger(x_1) \cdots \Psi^\dagger(x_{ln}) |0\rangle = 0. \quad (65)$$

It then follows by multiplication by

$$\langle 0| \Psi(x'_{ln}) \cdots \Psi(x'_1)\Psi(X'_n) \cdots \Psi(X'_1),$$

and use of Wick's theorem, that

$$\mathcal{S} \sum_{\alpha_1 \cdots \alpha_n} \Delta(\alpha_1 \cdots \alpha_n)\varphi_{\alpha_1}(X_1 x_1 \cdots x_l) \cdots \varphi_{\alpha_n}(X_n x_{ln-l+1} \cdots x_{ln}) \equiv 0, \quad (65)$$

where $\mathcal{S}$ antisymmetrizes its operand with respect to permutations of the x_i and symmetrizes or antisymmetrizes it with respect to permutations of the X_i, depending upon the nuclear statistics. But it follows from the fact that Δ satisfies (5) [or (7)] that the operand of $\mathcal{S}$ is already properly symmetrized and antisymmetrized; indeed, Eqs. (5) were derived as the necessary and sufficient conditions that an expansion of the form (2) be properly symmetric and antisymmetric under exchange of nuclei and electrons between atoms (symmetry and antisymmetry under exchanges *within* an atom is, of course, automatic). Hence (65) reduces simply to

$$\sum_{\alpha_1 \cdots \alpha_n} \Delta(\alpha_1 \cdots \alpha_n)\varphi_{\alpha_1}(X_1 x_1 \cdots x_l) \cdots \varphi_{\alpha_n}(X_n x_{ln-l+1} \cdots x_{ln}) \equiv 0. \quad (66)$$

It then follows from (1) that

$$\Delta(\alpha_1 \cdots \alpha_n) \equiv 0, \quad (67)$$

the desired result.

In summary: the problem of the redundancy of the physical atomic product states (53) is solved by the introduction of precisely the same subsidiary conditions (7) as were introduced in Sec. 2 for the purpose of ensuring the proper statistics under exchange of nuclei and electrons between different atoms; this is no accident, since the redundancy results precisely from the possibility of such exchanges.

In order to solve the problem of the complicated commutation relations between the A_α and the $A^\dagger_\beta$, we define, in analogy with Dyson's spin-wave treatment,[3] an ideal-atom state space $\mathfrak{IC}_{\text{ideal}}$. We first define ideal-atom annihilation and creation operators a_α and $a^\dagger_\alpha$ by their commutation relations

$$[a_\alpha, a_\beta]_- = [a^\dagger_\alpha, a^\dagger_\beta]_- = 0; \quad [a_\alpha, a^\dagger_\beta]_- = \delta_{\alpha\beta}, \quad (68)$$

together with the condition

$$a_\alpha |0) = 0; \quad (69)$$

where $|0)$ is the (normalized) ideal-atom vacuum state; commutators or anticommutators are to be taken in (68) according to whether the $A^\dagger_\alpha$ commute or anticommute, i.e., according to whether $2J + l$ is even or odd. Denote by $\mathfrak{IC}$ the physical many-atom state space consisting of the union, as n runs from zero to infinity, of all the n-nucleus, ln-electron spaces. Any state Φ in $\mathfrak{IC}$ can be represented in the form

$$\Phi = \sum_{n=0}^{\infty} \sum_{\alpha_1 \cdots \alpha_n} c_n(\alpha_1 \cdots \alpha_n) A^\dagger_{\alpha_1} \cdots A^\dagger_{\alpha_n} |0\rangle, \quad (70)$$

with coefficient functions c_n which are made unique by imposition of the subsidiary conditions (7). Given any Φ, we define its correspondent Φ_{ideal} by

$$\Phi_{\text{ideal}} = \sum_{n=0}^{\infty} \sum_{\alpha_1 \cdots \alpha_n} c_n(\alpha_1 \cdots \alpha_n) a^\dagger_{\alpha_1} \cdots a^\dagger_{\alpha_n} |0), \quad (71)$$

with the *same* coefficient functions c_n, and define the ideal-atom space $\mathfrak{IC}_{\text{ideal}}$ to be the set[17] of all such states Φ_{ideal} as Φ runs over all of $\mathfrak{IC}$. This establishes a one-one correspondence $\mathfrak{IC} \leftrightarrow \mathfrak{IC}_{\text{ideal}}$. The subsidiary conditions are essential here, since without them one state in $\mathfrak{IC}$ would have many images in $\mathfrak{IC}_{\text{ideal}}$ due to the fact that the expansion (70) is only made unique by imposition of the subsidiary conditions; thus $\mathfrak{IC}_{\text{ideal}}$ is *not* the space of *all* linear combinations of ideal-atom product states $a^\dagger_{\alpha_1} \cdots a^\dagger_{\alpha_n} |0)$. Indeed, the ideal-atom product states themselves are not contained in $\mathfrak{IC}_{\text{ideal}}$ (except

[17] That $\mathfrak{IC}_{\text{ideal}}$ is a vector space, and not merely a set, follows from the linearity of the subsidiary conditions (7), which implies that any linear combination of states of the form (71), with c_n's satisfying the subsidiary conditions, is also a state of the same type.

in special cases[18]) for $n \geq 2$. In this respect, our treatment differs from that of Dyson. Nevertheless, in actual calculations it is convenient, and usually essential, to consider initially the whole state space generated unrestrictedly from $|0)$ by the $a_\alpha^\dagger$, only imposing the subsidiary conditions in the sense of (30) or to a given order of perturbation theory.

The correspondence between operators O on $\mathfrak{C}$ and those O_{ideal} on $\mathfrak{C}_{\text{ideal}}$ is determined by expressing $O\Phi$ in the form (70), using (71), and identifying the result as some operator O_{ideal} acting on Φ_{ideal}. The procedure to be followed for any observable is illustrated by the case of the Hamiltonian H, which is expressed in terms of nucleus and electron creation and annihilation operators as

$$H = T + V,$$

$$T = \int dX\, \Psi^\dagger(X)T(X)\Psi(X) + \int dx\, \Psi^\dagger(x)T(x)\Psi(x),$$

$$V = \frac{1}{2}\int dX\, dX'\, \Psi^\dagger(X)\Psi^\dagger(X')V(XX')\Psi(X')\Psi(X) + \frac{1}{2}\int dx\, dx'\, \Psi^\dagger(x)\Psi^\dagger(x')V(xx')\Psi(x')\Psi(x) + \int dX\, dx\, \Psi^\dagger(X)\Psi^\dagger(x)V(Xx)\Psi(x)\Psi(X), \quad (72)$$

where the single-particle operators T and two-partiele operators V are the same as in (8). Since H conserves both electron and nucleus numbers, it is sufficient (because of linearity) to obtain an expression for $H\Phi$, with Φ given by (55), as a state of the same type as Φ. One can furthermore first consider the operation of H on a single typical atomic product state (53). Expressing $HA_{\alpha_1}^\dagger \cdots A_{\alpha_n}^\dagger$ in normally-ordered form with the aid of Wick's theorem[13] and noting that $\Psi(X)$ and $\Psi(x)$ annihilate the vacuum, one obtains an expression for $HA_{\alpha_1}^\dagger \cdots A_{\alpha_n}^\dagger |0\rangle$ involving only creation operators $\Psi^\dagger$ acting on the vacuum; upon expressing these in terms of the $A_\alpha^\dagger$ with the aid of (1), one obtains the desired expansion of $HA_{\alpha_1}^\dagger \cdots A_{\alpha_n}^\dagger |0\rangle$ in terms of atomic product states, which can finally be substituted into $H\Phi$.

Consider first the single-particle part, T, or H. Contracting annihilation operators in T with creation operators in $A_{\alpha_1}^\dagger \cdots A_{\alpha_n}^\dagger$ according to Wick's theorem, one finds by (72), (51), and the antisymmetry and symmetry of the φ_α that

$$TA_{\alpha_1}^\dagger \cdots A_{\alpha_n}^\dagger = \sum_{p=1}^{n} A_{\alpha_1}^\dagger \cdots A_{\alpha_{p-1}}^\dagger M_{\alpha_p} A_{\alpha_{p+1}}^\dagger \cdots A_{\alpha_n}^\dagger + (\text{terms with } \Psi \text{ operators on the right}), \quad (73)$$

where

$$M_\alpha = (l!)^{-\frac{1}{2}} \int dX\, dx_1 \cdots dx_l \times \Psi^\dagger(X)\Psi^\dagger(x_1)[T(X) + lT(x_1)] \times \varphi_\alpha(Xx_1 \cdots x_l)\Psi^\dagger(x_2) \cdots \Psi^\dagger(x_l). \quad (74)$$

Expanding $[T(X) + lT(x_1)]\varphi_\alpha(Xx_1 \cdots x_l)$ in terms of the $\varphi_\beta(Xx_1 \cdots x_l)$ and using (51), one finds

$$M_\alpha = \sum_\beta (\beta\, |T|\, \alpha) A_\beta^\dagger, \quad (75)$$

where $(\beta\, |T|\, \alpha)$ is defined by (11); hence,

$$TA_{\alpha_1}^\dagger \cdots A_{\alpha_n}^\dagger |0\rangle = \sum_\alpha \sum_{p=1}^{n} (\alpha\, |T|\, \alpha_p) A_{\alpha_1}^\dagger \cdots A_{\alpha_{p-1}}^\dagger \times A_\alpha^\dagger A_{\alpha_{p+1}}^\dagger \cdots A_{\alpha_n}^\dagger |0\rangle. \quad (76)$$

Consider next the two-particle part, V, of H. There are two different types of terms arising when one makes contractions of V with $A_{\alpha_1}^\dagger \cdots A_{\alpha_n}^\dagger$. The first type arise from contractions of both annihilation operators Ψ in a product $\Psi^\dagger\Psi^\dagger\Psi\Psi$ with creation operators $\Psi^\dagger$ in the *same* $A_\alpha^\dagger$ operator, and the second from contractions which couple two $A^\dagger$ operators, i.e. in which the two Ψ operators are contracted with $\Psi^\dagger$ operators in *different* $A^\dagger$ operators. The first type can be handled by a derivation completely parallel to that of (76); the second type is more complicated, but reducible by an obvious generalization of the simpler derivation. We give only the result:

$$VA_{\alpha_1}^\dagger \cdots A_{\alpha_n}^\dagger |0\rangle = (V_0 + V')A_{\alpha_1}^\dagger \cdots A_{\alpha_n}^\dagger |0\rangle,$$

$$V_0 A_{\alpha_1}^\dagger \cdots A_{\alpha_n}^\dagger |0\rangle = \sum_\alpha \sum_{p=1}^{n} (\alpha\, |V|\, \alpha_p) A_{\alpha_1}^\dagger \cdots A_{\alpha_{p-1}}^\dagger A_\alpha^\dagger A_{\alpha_{p+1}}^\dagger \cdots A_{\alpha_n}^\dagger |0\rangle,$$

$$V' A_{\alpha_1}^\dagger \cdots A_{\alpha_n}^\dagger |0\rangle = \sum_{\alpha\beta} \sum_{p<q}^{n} (\alpha\beta\, |V|\, \alpha_p\alpha_q) A_{\alpha_1}^\dagger \cdots A_{\alpha_{p-1}}^\dagger A_\alpha^\dagger A_{\alpha_{p+1}}^\dagger \cdots A_{\alpha_{q-1}}^\dagger A_\beta^\dagger A_{\alpha_{q+1}}^\dagger \cdots A_{\alpha_n}^\dagger |0\rangle, \quad (77)$$

where the matrix elements are given by (11). Substituting (76) and (77) into (55), one finds

$$H\Phi = \sum_{\alpha_1 \cdots \alpha_n} [Hc(\alpha_1 \cdots \alpha_n)] A_{\alpha_1}^\dagger \cdots A_{\alpha_n}^\dagger |0\rangle,$$

$$Hc(\alpha_1 \cdots \alpha_n) \equiv Tc(\alpha_1 \cdots \alpha_n) + V_0 c(\alpha_1 \cdots \alpha_n) + V'c(\alpha_1 \cdots \alpha_n), \quad (78)$$

[18] E.g., for a system of ("elementary") identical noninteracting bosons with periodic boundary conditions, the states $(a_0^\dagger)^n|0)$ satisfy the subsidiary condition for arbitrary n if one considers "atoms" consisting of two (noninteracting) bosons, since the ground state $\varphi_0(XX')$ is then just the constant Ω^{-1} where Ω is the quantization volume.

where T, V_0, and V' are defined as operators on the coefficients c by (10). Because of (58), there are many different expansions of $HA^\dagger_{\alpha_1} \cdots A^\dagger_{\alpha_n}|0\rangle$ in terms of the $A^\dagger_{\alpha_1'} \cdots A^\dagger_{\alpha_n'}|0\rangle$, the particular expansions (76) and (77) being distinguished only by their simplicity. Nevertheless, it follows from the subsidiary conditions (7) on c that Hc is the *unique* coefficient function for $H\Phi$ satisfying the subsidiary conditions [see discussion following Eq. (25)]. Since (78) holds for all n, one has then, by (71),

$$H_{\text{ideal}}\Phi_{\text{ideal}} = \sum_{n=0}^{\infty} [Hc_n(\alpha_1 \cdots \alpha_n)]a^\dagger_{\alpha_1} \cdots a^\dagger_{\alpha n}|0). \quad (79)$$

If one now asks what operator H_{ideal}, expressed as an explicit function of the a_α and $a^\dagger_\alpha$ operators, leads to (79) with the expressions (10) for Tc, V_0c, and $V'c$, one finds precisely the expression (20) of Sec. 2. One can similarly show that the subsidiary conditions can be expressed in $\mathfrak{K}_{\text{ideal}}$ in the form (25) with the exchange operators given by (24). Finally, the total atom number operator[19]

$$N_{\text{ideal}} \equiv \int dX\, \Psi^\dagger(X)\Psi(X) = l^{-1}\int dx\, \Psi^\dagger(x)\Psi(x) \quad (80)$$

in $\mathfrak{K}$ transforms into $\mathfrak{K}_{\text{ideal}}$ as $\sum_\alpha a^\dagger_\alpha a_\alpha$, the same expression as given in Sec. 2 [Eq. (26)]. Because of the linear independence (in fact, orthogonality) of the *ideal* atomic product states $a^\dagger_{\alpha_1} \cdots a^\dagger_{\alpha_n}|0)$, these operators in $\mathfrak{K}_{\text{ideal}}$ are *uniquely* determined by their correspondents in $\mathfrak{K}$; the subsidiary conditions on $c(\alpha_1 \cdots \alpha_n)$ are essential for this uniqueness.

If L is any observable, then the eigenvalue equation

$$L\Phi = \lambda\Phi \quad (81)$$

in $\mathfrak{K}$ transforms into $\mathfrak{K}_{\text{ideal}}$ as

$$L_{\text{ideal}}\Phi_{\text{ideal}} = \lambda\Phi_{\text{ideal}}, \quad (82)$$

where Φ_{ideal} and L_{ideal} are the *unique* images in $\mathfrak{K}_{\text{ideal}}$ of the state Φ and operator L in $\mathfrak{K}$, determined by the above procedures. In order to verify (82), we write (81) in the form

$$\sum_{n=0}^{\infty} \sum_{\alpha_1 \cdots \alpha_n} [Lc_n(\alpha_1 \cdots \alpha_n) - \lambda c_n(\alpha_1 \cdots \alpha_n)]A^\dagger_{\alpha_1} \cdots A^\dagger_{\alpha_n}|0\rangle = 0, \quad (83)$$

where L is defined as an operator on c by equations analogous to (78) and (10). Since c_n satisfies the subsidiary conditions, so does $Lc_n - \lambda$ [see discussion following Eq. (25)]. It then follows from (64)–(67) that

$$Lc_n(\alpha_1 \cdots \alpha_n) - \lambda c_n(\alpha_1 \cdots \alpha_n) \equiv 0, \quad (84)$$

hence

$$\sum_{n=0}^{\infty} \sum_{\alpha_1 \cdots \alpha_n} [Lc_n(\alpha_1 \cdots \alpha_n) - \lambda c_n(\alpha_1 \cdots \alpha_n)]a^\dagger_{\alpha_1} \cdots a^\dagger_{\alpha n}|0) = 0, \quad (85)$$

where the zero on the right side is, or course, the null vector in $\mathfrak{K}_{\text{ideal}}$; Eq. (85) is equivalent to (82). Again, the subsidiary conditions on c are quite essential for the validity of (82); without them (84) would not be a necessary consequence of (83).

In summary: the formalism of this section, based on an analogy with Dyson's treatment of spin-wave interactions, turns out to be completely equivalent to the less abstract formalism of Sec. 2. The subsidiary conditions (7) [or, as an eigenvalue equation in $\mathfrak{K}_{\text{ideal}}$, (25)], which were invoked in Sec. 2 to ensure that an expansion such as (2) have the correct symmetry properties under exchange of nuclei and electrons between different atoms, appear here in a new guise, as the necessary and sufficient conditions for a *unique* expansion (55) of a state of $\mathfrak{K}$ in terms of the *over*complete set of atomic product states (53). This is necessary for a *unique* transcription of states and operators into $\mathfrak{K}_{\text{ideal}}$, hence for a correct transcription of eigenvalue equations. The identity of the subsidiary conditions in Sec. 2 and in this section results from the fact that the redundancy of the physical atomic product states (53) is a direct result of the possibility of exchange of electrons and nuclei between different atoms.

4. APPLICATION TO SYSTEMS OF IDENTICAL PARTICLES

In this section we shall apply the general formalism to systems of identical particles, considering "atoms" composed of two identical particles. In addition to providing an illustration of the general formalism, this will enable us to make contact with related work of Blatt and Matsubara[4] and to obtain new results concerning Bose condensation of fermion pairs in a superconductor.

We consider a system of identical particles described in the usual second-quantization formalism by a Hamiltonian of the form

$$H = \sum_k \epsilon(k)b^\dagger_k b_k + \tfrac{1}{2}\sum_{kk'k''k'''} v(kk'k''k''')b^\dagger_k b^\dagger_{k'} b_{k''} b_{k'''}, \quad (86)$$

where b_k and $b^\dagger_k$ create particles in the single-particle states labeled by k, which will usually denote

[19] The correctness of this expression can be inferred from the fact that each atom contains one nucleus and l electrons, or more precisely, from the fact that $\mathfrak{K}$ is the union of all n-nucleus, ln-electron spaces for n from zero to infinity; the two given expressions for N are equal on this space $\mathfrak{K}$.

momentum in the Bose case or momentum and spin in the Fermi case, although we are not restricted to that particular choice of single-particle states. The interaction v is completely general except for the hermiticity and symmetry requirements

$$v^*(kk'k''k''') = v(k''k'''kk'),$$
$$v(kk'k''k''') = \pm v(k'kk''k''') = \pm v(kk'k'''k''), \quad (87)$$

the upper or lower sign being taken depending upon whether the individual particles are bosons or fermions. The "atomic" second-quantization formalism can be introduced by either the method of Sec. 2 or that of Sec. 3; we choose the more elementary and transparent method of Sec. 2. In order to apply this method one must first express H in a $2n$-particle momentum- (and spin-)space representation. Given any $2n$-particle state Φ in the quantized-field representation, the corresponding $2n$-particle momentum wavefunction $\psi(k_1 \cdots k_{2n})$ is defined by the expansion

$$\Phi = [(2n)!]^{-\frac{1}{2}} \sum_{k_1 \cdots k_{2n}} \psi(k_1 \cdots k_{2n}) b^\dagger_{k_1} \cdots b^\dagger_{k_{2n}} |0\rangle. \quad (88)$$

Operating on this equation with (86) and using Wick's theorem,[13] one finds that H is defined as an operator on wavefunctions ψ by

$$H\psi(k_1 \cdots k_{2n}) = \sum_{i=1}^{2n} \epsilon(k_i)\psi(k_1 \cdots k_{2n}) + \sum_{kk'} \sum_{j<l}^{2n} v(k_j k_l kk') \times \psi(k_1 \cdots k_{j-1} k k_{j+1} \cdots k_{l-1} k' k_{l+1} \cdots k_{2n}). \quad (89)$$

We next introduce any complete orthonormal set of two-particle momentum wavefunctions $\varphi_\alpha(kk')$ with the proper statistics, i.e. $\varphi_\alpha(k'k) = \pm\varphi_\alpha(kk')$, and expand ψ in terms of them:

$$\psi(k_1 \cdots k_{2n}) = \sum_{\alpha_1 \cdots \alpha_n} c(\alpha_1 \cdots \alpha_n)\varphi_{\alpha_1}(k_1k_2) \times \varphi_{\alpha_2}(k_3k_4) \cdots \varphi_{\alpha_n}(k_{2n-1}k_{2n}). \quad (90)$$

The problem is now to transform to a second-quantization representation in terms of *pair* annihilation and creation operators a_α, $a^\dagger_\alpha$ satisfying Bose statistics:

$$[a_\alpha, a_\beta] = [a^\dagger_\alpha, a^\dagger_\beta] = 0, \quad [a_\alpha, a^\dagger_\beta] = \delta_{\alpha\beta}. \quad (91)$$

The derivation proceeds exactly as in Sec. 2, by first expressing H as an operator on the wavefunctions $c(\alpha_1 \cdots \alpha_n)$ by expanding (89) in terms of the φ_α, and then introducing second quantization in the usual way. The final expression for the Hamiltonian is of the form (20) with

$$(\alpha\,|T|\,\beta) = 2\sum_{k_1k_2} \varphi^*_\alpha(k_1k_2)\epsilon(k_1)\varphi_\beta(k_1k_2),$$
$$(\alpha\,|V|\,\beta) = \sum_{k_1k_2}\sum_{k_1'k_2'} \varphi^*_\alpha(k_1k_2)v(k_1k_2k_1'k_2')\varphi_\beta(k_1'k_2'),$$
$$(\alpha\beta\,|V|\,\gamma\delta) = 4\sum_{k_1k_2}\sum_{k_1'k_2'}\sum_{k_1''k_2''} \varphi^*_\alpha(k_1k_2)\varphi^*_\beta(k_1'k_2') \times v(k_1k_1'k_1''k_2'')\varphi_\gamma(k_1''k_2)\varphi_\delta(k_2''k_2'). \quad (92)$$

The subsidiary condition analogous to (25) imposed on every allowable n-pair state Φ is

$$I\Phi = \pm\tfrac{1}{2}n(n-1)\Phi, \quad (93)$$

where the plus or minus sign is taken depending on whether the particles described by (86) are bosons or fermions, and where the exchange operator I is

$$I = \tfrac{1}{2}\sum_{\alpha\beta\gamma\delta} (\alpha\beta\,|I|\,\gamma\delta)a^\dagger_\alpha a^\dagger_\beta a_\delta a_\gamma, \quad (94)$$

with

$$(\alpha\beta\,|I|\,\gamma\delta) = \sum_{k_1k_2}\sum_{k_1'k_2'} \varphi^*_\alpha(k_1k_2)\varphi^*_\beta(k_1'k_2') \times \varphi_\gamma(k_1'k_2)\varphi_\delta(k_1k_2'). \quad (95)$$

A second-quantization formalism for fermion pairs has been derived previously, using Dyson's method of the ideal state space, by Blatt and Matsubara.[4] However, these authors did not specify any definite choice singling out one of the many different expansions of a given many-fermion state in terms of the redundant set of fermion-pair product states (see Sec. 3), i.e. they did not impose any subsidiary condition. As a result, their definition (4.10) of operators in the ideal state space is ambiguous. Dyson[3] avoids the corresponding difficulty of the redundancy of physical spin-wave product states by using a formula for the partition function which takes the redundancy into account, and an explicit construction of the Hamiltonian in the ideal spin-wave space which singles out one (no doubt the simplest) of the many possible images of the physical Hamiltonian. The problem of redundancy is solved in our treatment by the subsidiary condition (93), which, had we used the method of Sec. 3, would be the condition making the correspondence between states and operators in the physical and ideal state spaces one-one. In the method of Sec. 2 the subsidiary condition merely ensures that the wavefunction $\psi(k_1 \cdots k_{2n})$ [Eq. (90)] be symmetric or antisymmetric under particle exchanges between different pair wavefunctions φ_α, but, as shown in Sec. 3, this is equivalent to the more abstract interpretation of Sec. 3.

The simplest possible illustrations of the general formalism are given by the ideal Bose and Fermi

gases, for which $v \equiv 0$ in (86). The many-pair Hamiltonian is then simply

$$H = \sum_{\alpha\beta} (\alpha \,|T|\, \beta) a_\alpha^\dagger a_\beta. \tag{96}$$

In the Bose case the single-particle variables k are to be interpreted as momenta $\mathbf{k}$ quantized according to periodic boundary conditions, i.e. all components integral multiples of $2\pi\Omega^{-\frac{1}{3}}$ where Ω is the volume of the system. The noninteracting pair wavefunctions $\varphi_\alpha(\mathbf{kk}')$ in this case are simply

$$\varphi_{\mathbf{qq}'}(\mathbf{kk}') = [2(1 + \delta_{\mathbf{qq}'})]^{-\frac{1}{2}}(\delta_{\mathbf{kq}}\delta_{\mathbf{k}'\mathbf{q}'} + \delta_{\mathbf{kq}'}\delta_{\mathbf{k}'\mathbf{q}}), \tag{97}$$

where the two-particle quantum numbers α are pairs $(\mathbf{qq}')$ of single-particle momenta; note that the states $\varphi_{\mathbf{qq}'}$ and $\varphi_{\mathbf{q}'\mathbf{q}}$ are identical, so that care must be taken in sums over α, β, etc. to avoid overcounting. One finds

$$\begin{aligned}(\mathbf{q}_1\mathbf{q}_2 \,|T|\, \mathbf{q}_1'\mathbf{q}_2') = {} & [\epsilon(\mathbf{q}_1) + \epsilon(\mathbf{q}_2)] \\ & \times [(1 + \delta_{\mathbf{q}_1\mathbf{q}_2})(1 + \delta_{\mathbf{q}_1'\mathbf{q}_2'})]^{-\frac{1}{2}} \\ & \times (\delta_{\mathbf{q}_1\mathbf{q}_1'}\delta_{\mathbf{q}_2\mathbf{q}_2'} + \delta_{\mathbf{q}_1\mathbf{q}_2'}\delta_{\mathbf{q}_2\mathbf{q}_1'}),\end{aligned} \tag{98}$$

so that (96) becomes simply

$$H = \sum_{(\mathbf{q}_1\mathbf{q}_2)} [\epsilon(\mathbf{q}_1) + \epsilon(\mathbf{q}_2)] a_{\mathbf{q}_1\mathbf{q}_2}^\dagger a_{\mathbf{q}_1\mathbf{q}_2} \tag{99}$$

where

$$\sum_{(\mathbf{q}_1\mathbf{q}_2)} f(\mathbf{q}_1\mathbf{q}_2) \equiv \tfrac{1}{2} \sum_{\mathbf{q}_1\mathbf{q}_2}' f(\mathbf{q}_1\mathbf{q}_2) + \sum_{\mathbf{q}_1} f(\mathbf{q}_1\mathbf{q}_1), \tag{100}$$

f being any symmetric function of $\mathbf{q}_1$ and $\mathbf{q}_2$. Similarly, one finds for (93)

$$\begin{aligned}I = {} & \tfrac{1}{2} \sum_{(\mathbf{q}_1\mathbf{q}_2)} \sum_{(\mathbf{q}_1'\mathbf{q}_2')} \left[\frac{(1 + \delta_{\mathbf{q}_1'\mathbf{q}_2})(1 + \delta_{\mathbf{q}_1\mathbf{q}_2'})}{(1 + \delta_{\mathbf{q}_1\mathbf{q}_2})(1 + \delta_{\mathbf{q}_1'\mathbf{q}_2'})}\right]^{\frac{1}{2}} \\ & \times a_{\mathbf{q}_1'\mathbf{q}_2}^\dagger a_{\mathbf{q}_1\mathbf{q}_2'}^\dagger a_{\mathbf{q}_1'\mathbf{q}_2'} a_{\mathbf{q}_1\mathbf{q}_2}.\end{aligned} \tag{101}$$

The n-pair ground state of the Hamiltonian (99) is clearly

$$\Phi_0 = (n!)^{-\frac{1}{2}}(a_{00}^\dagger)^n \,|0), \tag{102}$$

since the $a_{\mathbf{qq}'}$ and $a_{\mathbf{qq}'}^\dagger$ operators satisfy Bose statistics (91); we are assuming that $\epsilon(\mathbf{q})$ is a minimum for $\mathbf{q} = 0$, as is the case for free particles with $\epsilon(\mathbf{q}) = \hbar^2q^2/2m$. In order that Φ_0 be an acceptable many-atom state it must satisfy the subsidiary condition (93) (with the plus sign). When I acts on Φ_0, one or the other of the two annihilation operators commutes through $(a_{00}^\dagger)^n$ and annihilates the vacuum unless $\mathbf{q}_1 = \mathbf{q}_2 = \mathbf{q}_1' = \mathbf{q}_2' = 0$. Thus one has simply

$$\begin{aligned}I\Phi_0 &= \tfrac{1}{2}(a_{00}^\dagger)^2 a_{00}^2 \Phi_0 \\ &= \tfrac{1}{2}N_{00}(N_{00} - 1)\Phi_0 = \tfrac{1}{2}n(n - 1)\Phi_0,\end{aligned} \tag{103}$$

the desired result; here $N_{00} = a_{00}^\dagger a_{00}$, the zero-momentum-pair occupation number operator. The ground-state energy is clearly $2n\epsilon(0)$, the well-known result for an ideal Bose gas of $2n$ particles (n pairs). On the other hand, the excited independent-pair product states do not in general[20] satisfy the subsidiary condition, although they are eigenstates of H. In order to satisfy the subsidiary condition for excited states one would have to form appropriate linear combinations of degenerate states[21] by considering, along with a given independent-pair state, all other states formed by interchanging particles (indices $\mathbf{q}$ or $\mathbf{q}'$ of $a_{\mathbf{qq}'}^\dagger$ operators) between different pairs. It is clear from (101) that I acting on such a linear combination gives another linear combination of the same type; the coefficients are to be chosen so that one in fact obtains an eigenstate of I with eigenvalue $\frac{1}{2}n(n - 1)$.

Consider next the ideal Fermi gas. The single-particle variables k in (92) and (95) are then to be interpreted as both momenta and spins, i.e. $k = (\mathbf{k}, \sigma)$. The pair wave functions $\varphi_\alpha(kk')$ are

$$\varphi_{qq'}(kk') = 2^{-\frac{1}{2}}(\delta_{kq}\delta_{k'q'} - \delta_{kq'}\delta_{k'q}), \tag{104}$$

where each q index also denotes both a momentum and a spin. Equation (99) is essentially unchanged; one has

$$H = \sum_{(q_1q_2)} [\epsilon(\mathbf{q}_1) + \epsilon(\mathbf{q}_2)] a_{q_1q_2}^\dagger a_{q_1q_2}, \tag{105}$$

where (q_1q_2) denotes an *ordered* pair with $q_1 \neq q_2$. Actual calculations are simplified if one writes

$$\sum_{(qq')} = \tfrac{1}{2} \sum_{qq'}', \tag{106}$$

and interprets[22]

$$a_{qq'} = -a_{q'q}, \tag{107}$$

since $\varphi_{\mathbf{qq}'}$ is antisymmetric in q and q'. The analogue of (101) is then

$$I = \tfrac{1}{2} \sum_{(q_1q_2)} \sum_{(q_1'q_2')} a_{q_1'q_2}^\dagger a_{q_1q_2'}^\dagger a_{q_1'q_2'} a_{q_1q_2}. \tag{108}$$

Since the $a_{qq'}$ and $a_{qq'}^\dagger$ operators satisfy Bose commutation relations, the n-pair ground state of H

[20] An exception occurs for the states

$$\Phi_{\mathbf{q}0} = [(n - 1)!]^{-\frac{1}{2}}(a_{00}^\dagger)^{n-1}a_{\mathbf{q}0}^\dagger|0)$$

with $\mathbf{q} \neq 0$, which do satisfy the subsidiary condition.

[21] E.g., neither $(a_{00}^\dagger)^{n-1}a_{\mathbf{qq}'}^\dagger|0)$ nor $(a_{00}^\dagger)^{n-2}a_{\mathbf{q}0}^\dagger a_{\mathbf{q}'0}^\dagger|0)$ satisfies the subsidiary condition, but one can form a linear combination of these two degenerate states which does satisfy the subsidiary condition.

[22] A useful corollary of (107) is that $a_{qq} = 0$.

is that in which all n pairs are Bose-condensed into the lowest pair state

$$\varphi_0(kk') = 2^{-\frac{1}{2}}\delta_{\mathbf{k}0}\delta_{\mathbf{k}'0}(\delta_{\delta\uparrow}\delta_{\delta'\downarrow} - \delta_{\delta\downarrow}\delta_{\delta'\uparrow}), \quad (109)$$

in which the fermions both have momentum zero, but opposite spins. Denoting the corresponding pair creation operator by $a_0^\dagger$, one sees that the n-pair ground state of H is

$$(n!)^{-\frac{1}{2}}(a_0^\dagger)^n\,|0). \quad (110)$$

But, in contradistinction to the case of the ideal Bose gas, (110) is not an allowable state of the ideal Fermi gas, because it does not satisfy the subsidiary condition (93) (with the *minus* sign).[23] Of course, we know that the correct ground state Φ_0 must, in fact, be the usual one in which each allowed momentum site within the Fermi sphere contains only two fermions of opposite spin (we assume that each fermion has total spin $\frac{1}{2}$). However, *before* application of the subsidiary condition, there is a spurious degeneracy associated with the many different ways of pairing the particles within the Fermi sphere. This degeneracy is removed by the subsidiary condition; in order to obtain an eigenstate of I one has to form a linear combination of states with all possible pairings. Let us define the Fermi sphere $\mathcal{S}_F$ to be the set of all $k = (\mathbf{k}, \sigma)$ for which $\mathbf{k}$ is an allowed (by periodic boundary conditions) momentum satisfying $|\mathbf{k}| < k_F$. For given k_F, this determines n, the number of allowed momentum sites within the Fermi sphere.[24] Denoting the $2n$ k values contained in $\mathcal{S}_F$, ordered in some arbitrary but well-defined way, by $k_1 \cdots k_{2n}$, we then define

$$\Phi_0 = [(2n)!]^{-\frac{1}{2}} \sum_P (-1)^{p(P)} \times Pa_{k_1k_2}^\dagger a_{k_3k_4}^\dagger \cdots a_{k_{2n-1}k_{2n}}^\dagger\,|0), \quad (111)$$

where $\sum_P$ is a sum over all permutations P of $k_1 \cdots k_{2n}$, $p(P)$ being the parity of P. Since the state (111) is completely antisymmetric in $k_1 \cdots k_{2n}$, and the subsidiary condition (93) (with the minus sign) is equivalent to such complete antisymmetry, it is clear that (111) must be an eigenstate of I with eigenvalue $-\frac{1}{2}n\,(n-1)$; this can be verified directly from (106)–(108).[25] It can furthermore be shown that (111) is the *lowest* such state; any state lower than (111) must have multiple occupation of some of the allowed k values within $\mathcal{S}_F$, and hence cannot be *completely* antisymmetrized as would be necessary to obtain an I eigenvalue of $-\frac{1}{2}n\,(n-1)$. The excited states can be treated in a similar fashion; again, the subsidiary condition is completely equivalent to the Pauli exclusion principle, although multiple occupation would be allowed by the *Bose* commutation relations of the $a_{kk'}^\dagger$ operators if the subsidiary condition were not imposed. The combinatorial problems associated with an exact evaluation of the generalized grand partition function

$$\Xi = \mathrm{Tr}\,\exp\,[-\beta(H - \mu N + \gamma I)] \quad (112)$$

[cf. (28)] to leading order in Ω are forbiddingly difficult for either the ideal Fermi or Bose gas, but on the basis of the general discussion following (31), such an evaluation should lead to the usual thermodynamic functions. This difficulty is somewhat academic, since as soon as a real interaction V' is present in H, it is necessary to introduce approximations, which should also be applicable to the exchange interaction γI. The treatment of the subsidiary condition with the aid of (112) and the analogue of (30) will then be justified for a given system and given approximation method if one can show that the fractional fluctuations of N and I vanish in the limit $n \to \infty$ when calculated to the given approximation.

The analysis of the previous paragraph shows that Bose condensation of fermion pairs is forbidden by the subsidiary condition in the case of the ideal Fermi gas, in spite of the fact that the pair annihilation and creation operators $a_{kk'}$ and $a_{kk'}^\dagger$ satisfy *Bose* commutation relations. What about an *interacting* Fermi gas? It has long been suspected that superconductivity is related to Bose condensation

[23] In fact, it is not difficult to show with the aid of (106)–(108), or directly from (94) and (95), that the state (110) is an eigenstate of I with eigenvalue $+\frac{1}{4}n(n-1)$. This is less than the eigenvalue $+\frac{1}{2}n(n-1)$ for the ideal Bose gas because of the *partial* antisymmetry implied by (107).

[24] This procedure ensures that the ground state be nondegenerate. On the other hand, for *given* n there does not always exist a k_F for which there are *exactly* n allowed momenta with $|\mathbf{k}| < k_F$; in such a case there is a *real* degeneracy of the ground state with respect to the spins and angular coordinates of the holes just under the surface of the Fermi sea. We wish to avoid the irrelevant complications of such a degeneracy.

[25] Using (106)–(108), the Bose commutation relations for the $a_{kk'}$ and $a_{kk'}^\dagger$ operators, and $a_{kk'}|0) = 0$, one can show that $Ia_{k_1k_2}^\dagger \cdots a_{k_{2n-1}k_{2n}}^\dagger|0) = \frac{1}{4}\sum_T' Ta_{k_1k_2}^\dagger \cdots a_{k_{2n-1}k_{2n}}^\dagger|0)$ where $\sum_T'$ is a sum over all interchanges T of two k indices on *different* $a^\dagger$ operators. The same relationship is valid if an arbitrary permutation is first applied to the k_j on both sides. But it follows from a standard argument of group theory that, for fixed T, the mapping $P \to TP$ is a *one-one* mapping of the set of all P onto itself. Furthermore, the parity of TP is the opposite of that of P, for any interchange T; hence, $\sum_P (-1)^{p(P)}P = -\sum_P (-1)^{p(P)}TP$. Finally, $T^2 = 1$ for any interchange T. Thus,

$$\begin{aligned} I\Phi_0 &= \tfrac{1}{4}[(2n)!]^{-\frac{1}{2}}\textstyle\sum_T' T \sum_P (-1)^{p(P)}Pa_{k_1k_2}^\dagger \cdots a_{k_{2n-1}k_{2n}}^\dagger|0) \\ &= -\tfrac{1}{4}[(2n)!]^{-\frac{1}{2}}\textstyle\sum_T' \sum_P (-1)^{p(P)}Pa_{k_1k_2}^\dagger \cdots a_{k_{2n-1}k_{2n}}^\dagger|0) \\ &= -\tfrac{1}{4}\nu(n)\Phi_0, \end{aligned}$$

where $\nu(n)$ is the number of different interchanges of two k indices on different $a^\dagger$ operators. One can readily show that $\nu(n) = 2n(n-1)$; hence $I\Phi_0 = -\frac{1}{2}n(n-1)\Phi_0$, the desired result.

of electron pairs,[26,27] and a detailed mathematical theory which incorporates such a condensation has been developed by Schafroth, Butler, Blatt, and Matsubara.[28,4,29] Furthermore, Blatt has given a simple qualitative argument[30] to the effect that Fermi statistics not only do not interfere with Bose condensation of fermion pairs, but actually aid it. These conclusions are, however, open to doubt. Finally, Blatt's qualitative argument[30] cannot be *generally* true, since it takes no account of interactions and hence should apply to an ideal Fermi gas, in contradiction with the fact that free fermion pairs do *not* condense. We shall therefore re-examine here the question of Bose condensation of fermion pairs for an interacting system; our approach will be to try to establish conditions under which such a condensation is or is not compatible with the subsidiary condition, i.e. with the exclusion principle.

Consider a two-fermion wavefunction $\varphi_0(kk')$ which is arbitrary except for the requirement that it is normalized, antisymmetric in k and k', and an eigenstate of total linear momentum and total S_z. The latter requirement implies that

$$\varphi_0(kk') = g_0(k)\delta_{k+k',q}, \tag{113}$$

where $q \equiv (\mathbf{q}, \sigma)$; the normalization and antisymmetry requirements are then

$$\sum_k |g_0(k)|^2 = 1, \qquad g_0(q-k) = -g_0(k). \tag{114}$$

The $2n$-fermion state in which all n fermion pairs are Bose-condensed into the state φ_0 is then (110), where $a_0^\dagger$ is the creation operator for the state φ_0; the other φ_α necessary to complete the formalism are also arbitrary, except for the requirement that the set composed of φ_0 plus all φ_α with $\alpha \neq 0$ must be complete and orthonormal. But it is clear from (94) and (95) that the state (110) will be an eigenstate of I if and only if

$$(\alpha\beta\,|I|\,00) = \delta_{\alpha 0}\delta_{\beta 0}\,(00\,|I|\,00), \tag{115}$$

in which case the eigenvalue of I will be $\frac{1}{2}n\,(n-1) \times (00\,|I|\,00)$. But by (95) and (113),

$$(00\,|I|\,00) = \sum_k |g_0(k)|^4 > 0. \tag{116}$$

Hence a state of the form (110) can never satisfy the subsidiary condition (93) (with the minus sign). This shows that *complete* Bose condensation into *any* momentum- and spin-conserving two-particle state φ_0 is prevented by the exclusion principle.

This result bears directly on the physical interpretation of the BCS theory.[31] The BCS variational trial state can be written in the form[32]

$$\Psi = \{\prod_{\mathbf{k}\in\mathcal{S}} [(1-h_\mathbf{k})^{\frac{1}{2}} + h_\mathbf{k}^{\frac{1}{2}} b_{\mathbf{k}\uparrow}^\dagger b_{-\mathbf{k}\downarrow}^\dagger]\}\,|0\rangle, \tag{117}$$

where $\mathcal{S}$ is the shell[31]

$$\epsilon_F - \hbar\omega < \epsilon(\mathbf{k}) < \epsilon_F + \hbar\omega; \tag{118}$$

$h_\mathbf{k}(=h_{-\mathbf{k}} = h_\mathbf{k}^*)$ is the solution of the BCS integral equation, which satisfies

$$\begin{aligned} h_\mathbf{k} &= 1, \quad \epsilon(\mathbf{k}) < \epsilon_F - \hbar\omega, \\ 0 < h_\mathbf{k} &< 1, \quad \mathbf{k} \in \mathcal{S}, \\ h_\mathbf{k} &= 0, \quad \epsilon(\mathbf{k}) > \epsilon_F + \hbar\omega; \end{aligned} \tag{119}$$

the $b_{\mathbf{k}\sigma}^\dagger$ are defined by

$$\begin{aligned} b_{\mathbf{k}\delta}^\dagger &\equiv c_{\mathbf{k}\delta}, \quad \epsilon(\mathbf{k}) < \epsilon_F - \hbar\omega, \\ &\equiv c_{\mathbf{k}\delta}^\dagger, \quad \epsilon(\mathbf{k}) > \epsilon_F - \hbar\omega, \end{aligned} \tag{120}$$

with $c_{\mathbf{k}\sigma}$ and $c_{\mathbf{k}\sigma}^\dagger$ the electron annihilation and creation operators, and $|0\rangle$ in (117) is the vacuum of the redefined Fermi operators $b_{\mathbf{k}\sigma}$, $b_{\mathbf{k}\sigma}^\dagger$. The $2n$-electron projection of (117) is[32,33]

$$\Psi_{\mathrm{BCS}} = \text{const}\,\{\sum_{\mathbf{k}\in\mathcal{S}} [h_\mathbf{k}/(1-h_\mathbf{k})]^{\frac{1}{2}} b_{\mathbf{k}\uparrow}^\dagger b_{-\mathbf{k}\downarrow}^\dagger\}^n\,|0\rangle. \tag{121}$$

Defining[34]

$$\begin{aligned} A_0^\dagger &\equiv \mathcal{C} \sum_{\mathbf{k}\in\mathcal{S}} [h_\mathbf{k}/(1-h_\mathbf{k})]^{\frac{1}{2}} b_{\mathbf{k}\uparrow}^\dagger b_{-\mathbf{k}\downarrow}^\dagger, \\ \mathcal{C} &= [\sum_{\mathbf{k}\in\mathcal{S}} h_\mathbf{k}/(1-h_\mathbf{k})]^{-\frac{1}{2}}, \end{aligned} \tag{122}$$

one can write (121) in the form

$$\Psi_{\mathrm{BCS}} = \text{const}\,(A_0^\dagger)^n\,|0\rangle. \tag{123}$$

Blatt[32] interprets this as a state in which all electrons in the shell $\mathcal{S}$ are Bose-condensed into the same pair state, but this interpretation is untenable. In the first place, A_0 and $A_0^\dagger$ do not satisfy the Bose commutation relation, so that (123) does not *directly* imply Bose condensation. Furthermore, even after passing to the ideal state space in which the pair annihilation and creation operators a_α, $a_\alpha^\dagger$ do satisfy Bose commutation relations, if one simply

[26] V. L. Ginzburg, Usp. Fiz. Nauk. **48**, 25 (1952).
[27] M. R. Schafroth, Phys. Rev. **96**, 1442 (1954).
[28] M. R. Schafroth, S. T. Butler, and J. M. Blatt, Helv. Phys. Acta **30**, 93 (1957).
[29] T. Matsubara and J. M. Blatt, Progr. Theoret. Phys. (Kyoto) **23**, 451 (1960); J. M. Blatt, J. Australian Math. Soc. **1**, 465 (1960); J. M. Blatt (to be published); T. Matsubara and C. J. Thompson (to be published).
[30] J. M. Blatt, Progr. Theoret. Phys. (Kyoto) **27**, 1137 (1962).
[31] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).
[32] J. M. Blatt, Progr. Theoret. Phys. (Kyoto) **23**, 447 (1960).
[33] Note that $2n$ is not the total number of electrons but rather the number in the shell $\mathcal{S}$.
[34] The normalization constant $\mathcal{C}$ is chosen so that $\langle 0|A_0 A_0^\dagger|0\rangle = 1$ [cf. (52)].

replaces $A_0^\dagger$ by $a_0^\dagger$ [and $|0\rangle$ by the ideal vacuum $|0)$], one obtains a state of the form (110)[35] which, according to the previous paragraph, cannot satisfy the subsidiary condition (93).[36] One can project out a state which does satisfy the subsidiary condition with the aid of the projection operator

$$P = (2\pi)^{-1} \int_0^{2\pi} d\theta\, e^{i[I+\frac{1}{2}n(n-1)]\theta}, \qquad (124)$$

but since I [Eq. (94)] involves not only a_0 and $a_0^\dagger$ but annihilation and creation operators for a whole complete set of pair states, it is certainly not true that the projected state has *complete* Bose condensation into the pair state φ_0. Following Blatt, one can regard the BCS ground state as exhibiting complete condensation of a certain type, called "Schafroth condensation" by Blatt for historical reasons, but our analysis shows that complete Schafroth condensation is not the same as *complete* Bose condensation. It is, however, plausible that there can be *partial* Bose condensation into a *bound* pair state φ_0 such as that of the BCS theory, and that the condensed fraction will approach unity as the density approaches zero. This is best investigated by transforming the BCS Hamiltonian into the electron-pair–boson representation with the aid of (92) and then examining the structure of its ground state (more exactly, that of $H + \gamma I$) in this representation. It would take us too far afield to do this here; we hope to present such an analysis in a later publication.

Another type of application which we do not have space to discuss here is the treatment of systems of identical particles with interactions containing a hard core. We wish only to point out the following: If the wavefunction $\psi(\mathbf{r}_1 \cdots \mathbf{r}_{2n})$ of a system of $2n$ identical particles is expanded in terms of two-particle wavefunctions $\varphi_\alpha(\mathbf{r}\mathbf{r}')$ according to (90) (with k's replaced by $\mathbf{r}$'s) and if each φ_α vanishes for $|\mathbf{r} - \mathbf{r}'| \leq a$ (one could, e.g., choose the complete set of solutions of the two-body hard-sphere problem), then clearly $\psi(\mathbf{r}_1 \cdots \mathbf{r}_{2n})$ will vanish whenever $|\mathbf{r}_1 - \mathbf{r}_2| \leq a$ or $|\mathbf{r}_3 - \mathbf{r}_4| \leq a \cdots$. But if the coefficient function $c(\alpha_1 \cdots \alpha_n)$ satisfies the subsidiary condition, or equivalently if the state vector Φ satisfies the subsidiary condition (93), then ψ will be completely symmetric or antisymmetric in $\mathbf{r}_1 \cdots \mathbf{r}_{2n}$, so that ψ will also vanish when *any* two arguments are closer than a. In this way, the problem of making ψ vanish when hard cores overlap is replaced by that of making Φ satisfy the subsidiary condition ensuring Bose or Fermi statistics. This problem is still nontrivial, but it may well be simpler than that of avoiding hard-core energy divergences.

[35] The corresponding pair state φ_0 is $\varphi_0(\mathbf{k}\sigma, \mathbf{k}'\sigma') = 2^{-\frac{1}{2}}\mathcal{C}(\delta_{\sigma\uparrow}\delta_{\sigma'\downarrow} - \delta_{\sigma\downarrow}\delta_{\sigma'\uparrow})\delta_{\mathbf{k}'\sigma-\mathbf{k}}[h_\mathbf{k}/(1-h_\mathbf{k})]^{\frac{1}{2}}$ for $\mathbf{k}$ and $\mathbf{k}'$ in the shell $\mathcal{S}$, and vanishes outside $\mathcal{S}$.

[36] In the terminology of Sec. 3, the coefficient function $c(\alpha_1 \cdots \alpha_n) = \delta_{\alpha_1 0} \cdots \delta_{\alpha_n 0}$ does not satisfy the subsidiary condition.

ACKNOWLEDGMENTS

I am very much indebted to G. Wentzel for many stimulating discussions of this work. To him is due the very important observation that the atomic second-quantization formalism which I originally obtained by an abstract method (Sec. 3) suggested by that of Dyson[3] ought to be obtainable by a more elementary method; this provided the stimulus for Sec. 2. I thank J. M. Blatt for correspondence which helped to clarify the relationship between his work and mine.

APPENDIX. EIGENVALUE SPECTRA OF THE EXCHANGE OPERATORS

The analysis is simplest if we return to the Schrödinger representation. The variables which are not permuted need not be indicated explicitly; thus one has [cf. (4) and (7)]

$$I = \sum_{p<q}^{n} I_{pq},$$

$$I_{pq}\psi(y_1 \cdots y_n) \equiv \psi(y_1 \cdots y_q \cdots y_p \cdots y_n), \qquad (A1)$$

where I is either I_{nuc} or I_{elec}, and $y_1 \cdots y_n$ are either $X_1 \cdots X_n$ or $x_1 x_{l+1} \cdots x_{ln-l+1}$. Since any permutation can be written as a product of interchanges, the problem of the possible behaviors of ψ when acted upon by the interchanges I_{pq}, and hence by their sum I, is closely related to the well-known problem of the possible symmetry classes of the group of all permutations of n objects. The solution of this problem can be expressed in terms of Young diagrams and their associated Young symmetry operators[37]; the set of all irreducible representations of the permutation group is in one–one correspondence with the set of all Young diagrams, which in turn is in one–one correspondence with the set of all partitions of n of the form $n_1 + n_2 + \cdots = n$, where the n_i are positive integers satisfying $n_1 \geq n_2 \geq \cdots$. The Young symmetry operator $S(\mathfrak{D})$ associated with a given Young diagram $\mathfrak{D}$ symmetrizes ψ with respect to the arguments lying in each row of $\mathfrak{D}$, and antisymmetrizes it with respect

[37] See, e.g., H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover Publications, Inc., New York, 1950), pp. 358 ff.

to the variables in each column; the functions $S(\mathfrak{D})\psi$ and $S(\mathfrak{D}')\psi$ are orthogonal for $\mathfrak{D} \neq \mathfrak{D}'$ because they have different symmetry.[37] We shall now prove that $S(\mathfrak{D})\psi$ is an eigenstate of I with an eigenvalue $l(\mathfrak{D})$ given by

$$l(\mathfrak{D}) = \tfrac{1}{2}n_1(n_1 - 1) + \tfrac{1}{2}n_2(n_2 - 1) + \cdots - \tfrac{1}{2}n_1^*(n_1^* - 1) - \tfrac{1}{2}n_2^*(n_2^* - 1) - \cdots, \quad \text{(A2)}$$

where the n_i are the lengths of the rows of $\mathfrak{D}$ and the n_i^* are the lengths of its columns. We first write the symmetry operator in the form[38]

$$S(\mathfrak{D}) = \sum_P c(P)P, \quad \text{(A3)}$$

where the summation runs over all permutations of the n arguments of ψ. The coefficients $c(P)$ are defined as follows: If P can be factored in the form $P = CR$, where R permutes the arguments in each row of $\mathfrak{D}$ among themselves and C similarly permutes the arguments in each column, then this factorization is unique,[38] and we define $c(P) = \pm 1$ depending upon whether the column permutation C is even or odd; if P cannot be so factored, then $c(P)$ is defined to be zero. Let us examine the products $I_{pq}S(\mathfrak{D})$. Each I_{pq} is itself a permutation (a single interchange), and so $I_{pq}P$ is another P; furthermore, by a standard argument of group theory, the mapping $P \rightarrow I_{pq}P$ is one–one, i.e. it *permutes* the set $\{P\}$. Hence (A3) can be written

$$S(\mathfrak{D}) = \sum_P c(I_{pq}P)I_{pq}P. \quad \text{(A4)}$$

Then, since $I_{pq}^2 = 1$, one has

$$IS(\mathfrak{D}) = \sum_{p<q}^{n} I_{pq}S(\mathfrak{D}) = \sum_P d(P)P, \quad \text{(A5)}$$

with

$$d(P) = \sum_{p<q}^{n} c(I_{pq}P). \quad \text{(A6)}$$

According to a theorem of Weyl,[39] it would follow from the relationships

$$d(PR) = d(P), \qquad d(CP) = (-1)^{p(C)}\, d(P), \quad \text{(A7)}$$

where R and C are any permutations of the types defined after (A3) and $p(P)$ is the parity of P, that

$$d(P) = l(\mathfrak{D})c(P), \quad \text{(A8)}$$

where $l(\mathfrak{D})$ is some numerical coefficient; it would then follow that $S(\mathfrak{D})\psi$ is an eigenstate of I with eigenvalue $l(\mathfrak{D})$. In order to prove the first of Eqs. (A7), we first write, by (A6),

$$d(PR) = \sum_{p<q}^{n} c(I_{pq}PR). \quad \text{(A9)}$$

By definition $c(I_{pq}PR)$ vanishes unless there exists a (unique) factorization

$$I_{pq}PR = CR', \quad \text{(A10)}$$

in which case $c(I_{pq}PR) = \pm 1$ according to whether C is even or odd. But (A10) can be written

$$I_{pq}P = CR'R^{-1} = CR'', \quad \text{(A11)}$$

since the inverse of a row permutation is also a row permutation; hence

$$c(I_{pq}PR) = c(I_{pq}P). \quad \text{(A12)}$$

The first of Eqs. (A7) then follows from (A6) and (A9). To prove the second of Eqs. (A7), we note first that a permutation C which permutes arguments in the same *column* of $\mathfrak{D}$ permutes arguments in the same *row* of the dual diagram $\mathfrak{D}^*$ (that differing from $\mathfrak{D}$ by interchange of rows with columns); it should therefore be possible to derive the second (A7) by applying the first (A7) to the dual diagram. Denoting quantities pertaining to $\mathfrak{D}^*$ by stars, one has by the first (A7), together with (A6) and (A9),

$$d^*(P^{-1}R^*) = \sum_{p<q}^{n} c^*(I_{pq}P^{-1}R^*) = \sum_{p<q}^{n} c^*(I_{pq}P^{-1}) = d^*(P^{-1}). \quad \text{(A13)}$$

But[40]

$$c^*(P) = (-1)^{p(P)}c(P^{-1}). \quad \text{(A14)}$$

Hence, since $I_{pq}^{-1} = I_{pq}$,

$$\sum_{p<q}^{n} (-1)^{p(I_{pq}P^{-1}R^*)}c((R^*)^{-1}PI_{pq}) = \sum_{p<q}^{n} (-1)^{p(I_{pq}P^{-1})}c(PI_{pq}). \quad \text{(A15)}$$

But

$$(-1)^{p(I_{pq}P^{-1}R^*)} = (-1)^{p(I_{pq})}(-1)^{p(P^{-1})}(-1)^{p(R^*)} = -(-1)^{p(P^{-1})}(-1)^{p(R^*)},$$

$$(-1)^{p(I_{pq}P^{-1})} = (-1)^{p(I_{pq})}(-1)^{p(P^{-1})} = -(-1)^{p(P^{-1})}. \quad \text{(A16)}$$

Canceling the constant factor $-(-1)^{p(P^{-1})}$ from both sides of (A15), one has then

$$(-1)^{p(R^*)} \sum_{p<q}^{n} c((R^*)^{-1}PI_{pq}) = \sum_{p<q}^{n} c(PI_{pq}). \quad \text{(A17)}$$

[38] Reference 37, p. 361.
[39] Theorem (14.2) of reference 37.
[40] Reference 37, p. 368.

But R^* permutes rows of $\mathfrak{D}^*$, hence columns of $\mathfrak{D}$, so one can give $(R^*)^{-1}$ the new name C. Then, since $p(C^{-1}) = p(C)$,

$$\sum_{p<q}^{n} c(CPI_{pq}) = (-1)^{p(C)} \sum_{p<q}^{n} c(PI_{pq}). \quad \text{(A18)}$$

Now

$$PI_{pq} = (PI_{pq}P^{-1})P, \; CPI_{pq} = [CPI_{pq}(CP)^{-1}]CP. \quad \text{(A19)}$$

But for any P, the transformation $I_{pq} \rightarrow PI_{pq}P^{-1}$ is a one–one mapping of the set $\{I_{pq}\}$ onto itself. Hence

$$\sum_{p<q}^{n} c(I_{pq}CP) = (-1)^{p(C)} \sum_{p<q}^{n} c(I_{pq}P), \quad \text{(A20)}$$

although, in general, the individual (pq) terms on the left and the right sides of (A20) are not equal (those on the right are a permutation of those on the left). The second of Eqs. (A7) then follows from (A20) and the definition (A6). This completes the proof of (A7), and hence of (A8).

It then follows from (A5) and (A8), that

$$IS(\mathfrak{D})\psi = l(\mathfrak{D})S(\mathfrak{D})\psi, \quad \text{(A21)}$$

i.e. that $S(\mathfrak{D})\psi$ is an eigenstate of I with eigenvalue $l(\mathfrak{D})$. To obtain the explicit expression for $l(\mathfrak{D})$, we note that $c(1) = 1$ where 1 is the identity permutation; hence by (A8) and (A6),

$$l(\mathfrak{D}) = d(1) = \sum_{p<q}^{n} c(I_{pq}). \quad \text{(A22)}$$

Now it follows from the definition of c that $c(I_{pq})$ vanishes unless y_p and y_q are in the same row or the the same column of $\mathfrak{D}$, in which cases it is $+1$ or -1, respectively; Eq. (A2) then follows immediately. Finally, since any ψ can be written as a linear combination of the states $S(\mathfrak{D})\psi$ with all Young diagrams $\mathfrak{D}$, Eq. (A2) gives *all* eigenvalues of I for n-variable functions $\psi(y_1 \cdots y_n)$, and hence all eigenvalues of I_{nuc} and I_{elec} for n-atom states $\psi(X_1 \cdots X_n x_1 \cdots x_{ln})$. It follows from (A2) that the spectra of I_{nuc} and I_{elec} for such states are discrete and lie on the interval $[-\frac{1}{2}n(n-1), +\frac{1}{2}n(n-1)]$.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 8 AUGUST 1963

Evaluation of Certain Sums Arising in Chemical Physics*

PHILIP B. ABRAHAM AND GEORGE H. WEISS
University of Maryland, College Park, Maryland
(Received 30 April 1963)

In this paper we show that the Poisson transformation of finite sums of the form

$$\sum_{j=1}^{N} F(\cos 2\pi j/N) \cos 2\pi jn/N$$

generally leads to a quickly convergent sum for problems of physical interest. Examples related to molecular orbital theory and lattice dynamics are discussed.

THERE are many problems in chemical physics in which sums of the form

$$S_N(n) = \frac{1}{N} \sum_{j=1}^{N} F\left[\cos \frac{2\pi j}{N}\right] \cos \frac{2\pi jn}{N}, \quad (1)$$

or the generalization to a multiple sum, form a central role. In many instances N is so large that the approximation of the sum by an integral is sufficiently accurate. However, there are problems in which N is of the order of 20 or 30, making it awkward to work with the sum of Eq. (1) directly, and the limit $N = \infty$ is not sufficiently accurate. It is the purpose of this paper to indicate a convenient transformation of the sum in Eq. (1) which allows one to estimate correction terms for large N.

As a more concrete motivation for the theory that follows, we cite two problems in which sums of the form given in Eq. (1) arise. In a study of bond alternation in long polyenes, Salem obtains sums of the form

$$S_{2n+1}(h) = \frac{1}{2n+1} \times \sum_{j=-n}^{n} \frac{\cos[2\pi hj/(2n+1)]}{\{1 + 2t^2 \cos[2\pi j/(2n+1)] + t^4\}^{\frac{1}{2}}}, \quad (2)$$

which is evaluated to lowest order in t^2.[1] Similar

* This research was supported in part by the United States Office of Naval Research under grant Nonr 595(17).

[1] L. Salem, Proc. Cambridge Phil. Soc. 57, 353 (1961).

But R^* permutes rows of $\mathfrak{D}^*$, hence columns of $\mathfrak{D}$, so one can give $(R^*)^{-1}$ the new name C. Then, since $p(C^{-1}) = p(C)$,

$$\sum_{p<q}^{n} c(CPI_{pq}) = (-1)^{p(C)} \sum_{p<q}^{n} c(PI_{pq}). \quad \text{(A18)}$$

Now

$$PI_{pq} = (PI_{pq}P^{-1})P, \; CPI_{pq} = [CPI_{pq}(CP)^{-1}]CP. \quad \text{(A19)}$$

But for any P, the transformation $I_{pq} \rightarrow PI_{pq}P^{-1}$ is a one–one mapping of the set $\{I_{pq}\}$ onto itself. Hence

$$\sum_{p<q}^{n} c(I_{pq}CP) = (-1)^{p(C)} \sum_{p<q}^{n} c(I_{pq}P), \quad \text{(A20)}$$

although, in general, the individual (pq) terms on the left and the right sides of (A20) are not equal (those on the right are a permutation of those on the left). The second of Eqs. (A7) then follows from (A20) and the definition (A6). This completes the proof of (A7), and hence of (A8).

It then follows from (A5) and (A8), that

$$IS(\mathfrak{D})\psi = l(\mathfrak{D})S(\mathfrak{D})\psi, \quad \text{(A21)}$$

i.e. that $S(\mathfrak{D})\psi$ is an eigenstate of I with eigenvalue $l(\mathfrak{D})$. To obtain the explicit expression for $l(\mathfrak{D})$, we note that $c(1) = 1$ where 1 is the identity permutation; hence by (A8) and (A6),

$$l(\mathfrak{D}) = d(1) = \sum_{p<q}^{n} c(I_{pq}). \quad \text{(A22)}$$

Now it follows from the definition of c that $c(I_{pq})$ vanishes unless y_p and y_q are in the same row or the the same column of $\mathfrak{D}$, in which cases it is $+1$ or -1, respectively; Eq. (A2) then follows immediately. Finally, since any ψ can be written as a linear combination of the states $S(\mathfrak{D})\psi$ with all Young diagrams $\mathfrak{D}$, Eq. (A2) gives *all* eigenvalues of I for n-variable functions $\psi(y_1 \cdots y_n)$, and hence all eigenvalues of I_{nuc} and I_{elec} for n-atom states $\psi(X_1 \cdots X_n x_1 \cdots x_{ln})$. It follows from (A2) that the spectra of I_{nuc} and I_{elec} for such states are discrete and lie on the interval $[-\frac{1}{2}n(n-1), +\frac{1}{2}n(n-1)]$.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 8 AUGUST 1963

Evaluation of Certain Sums Arising in Chemical Physics*

PHILIP B. ABRAHAM AND GEORGE H. WEISS
University of Maryland, College Park, Maryland
(Received 30 April 1963)

In this paper we show that the Poisson transformation of finite sums of the form

$$\sum_{j=1}^{N} F(\cos 2\pi j/N) \cos 2\pi jn/N$$

generally leads to a quickly convergent sum for problems of physical interest. Examples related to molecular orbital theory and lattice dynamics are discussed.

THERE are many problems in chemical physics in which sums of the form

$$S_N(n) = \frac{1}{N} \sum_{j=1}^{N} F\left[\cos \frac{2\pi j}{N}\right] \cos \frac{2\pi jn}{N}, \quad (1)$$

or the generalization to a multiple sum, form a central role. In many instances N is so large that the approximation of the sum by an integral is sufficiently accurate. However, there are problems in which N is of the order of 20 or 30, making it awkward to work with the sum of Eq. (1) directly, and the limit $N = \infty$ is not sufficiently accurate. It is the purpose of this paper to indicate a convenient transformation of the sum in Eq. (1) which allows one to estimate correction terms for large N.

As a more concrete motivation for the theory that follows, we cite two problems in which sums of the form given in Eq. (1) arise. In a study of bond alternation in long polyenes, Salem obtains sums of the form

$$S_{2n+1}(h) = \frac{1}{2n+1} \times \sum_{j=-n}^{n} \frac{\cos[2\pi hj/(2n+1)]}{\{1 + 2t^2 \cos[2\pi j/(2n+1)] + t^4\}^{\frac{1}{2}}}, \quad (2)$$

which is evaluated to lowest order in t^2.[1] Similar

* This research was supported in part by the United States Office of Naval Research under grant Nonr 595(17).

[1] L. Salem, Proc. Cambridge Phil. Soc. 57, 353 (1961).

sums arise in many calculations which use molecular orbital theory as a starting point. Another application of the present work is to the calculation of Green's functions for problems in lattice dynamics. If cyclic boundary conditions are assumed, the Green's function for a three-dimensional simple cubic lattice with nearest-neighbor interactions only, is[2]

$$g(m_1, m_2, m_3; z) = \frac{1}{N^3} \sum_{j_1=1}^{N} \sum_{j_2=1}^{N} \sum_{j_3=1}^{N} \frac{\cos \frac{2\pi j_1 m_1}{N} \cos \frac{2\pi j_2 m_2}{N} \cos \frac{2\pi j_3 m_3}{N}}{z + \alpha_1 \cos \frac{2\pi j_1}{N} + \alpha_2 \cos \frac{2\pi j_2}{N} + \alpha_3 \cos \frac{2\pi j_3}{N}}. \tag{3}$$

The method which we propose to elaborate is a form of the Poisson transformation, and is related to a technique exploited by Salem, although he did not take full advantage of it. Let us suppose that the function $F(\cos\theta)$ has a convergent Fourier expansion which can be written

$$F(\cos\theta) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos n\theta, \tag{4}$$

where

$$a_n = \frac{1}{\pi} \int_0^{2\pi} F(\cos\theta) \cos n\theta \, d\theta, \tag{5}$$

and that when this is substituted into Eq. (1) the summations over j and n can be interchanged. This will be the case, for example, when $\sum_n a_n$ is uniformly and absolutely convergent; a condition usually fulfilled in sums arising from physical problems. Then the expression for $S_N(r)$ is

$$\begin{aligned} S_N(r) &= \frac{1}{N} \sum_{j=1}^{N} \left[\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos \frac{2\pi jn}{N} \right] \cos \frac{2\pi jr}{N} \\ &= \frac{1}{2} \sum_{j=0}^{\infty} a_{r+jN} + \frac{1}{2} \sum_{j=1}^{\infty} a_{jN-r} \\ &= \frac{1}{2} \sum_{j=-\infty}^{\infty} a_{|jN+r|}. \end{aligned} \tag{6}$$

Thus the finite sum $S_N(r)$ is converted into an infinite sum of Fourier coefficients. While this may not seem to be helpful, it must be remembered that Fourier coefficients decrease with increasing index, by virtue of the Riemann–Lebesgue lemma, and the convergence of the resulting series can be quite rapid. In many cases it is sufficient to approximate $S_N(r)$ by the lowest-order terms

$$S_N(r) \sim \tfrac{1}{2}a_r + \tfrac{1}{2}a_{N-r}. \tag{7}$$

The term a_r is just the approximation of the sum by an integral, and $\frac{1}{2}a_{N-r}$ is the first correction term.

As a first example of our technique, let us evaluate Salem's sum in Eq. (2). In order to evaluate the sum we need the Fourier coefficients

$$\begin{aligned} a_j &= \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{\cos j\theta \, d\theta}{(1 + 2t^2 \cos\theta + t^4)^{\frac{1}{2}}} \\ &= \frac{2\pi^{\frac{1}{2}}}{\Gamma(j + \frac{1}{2})} \cdot \frac{1}{(1 - t^4)^{\frac{1}{2}}} P^j_{-\frac{1}{2}}\left(\frac{1 + t^4}{1 - t^4}\right), \end{aligned} \tag{8}$$

where $P^j_{-\frac{1}{2}}(x)$ is a generalized Legendre function of the first kind. Hence the formal expression for $S_{2n+1}(h)$ is

$$\begin{aligned} S_{2n+1}(h) &= \frac{1}{(1 - t^4)^{\frac{1}{2}}} P_{-\frac{1}{2}}\left(\frac{1 + t^4}{1 - t^4}\right) \delta_{h,0} \\ &+ \frac{\pi^{\frac{1}{2}}}{(1 - t^4)^{\frac{1}{2}}} \left[\sum_{j=1}^{\infty} \frac{1}{\Gamma(jN + h + \frac{1}{2})} P^{jN+h}_{-\frac{1}{2}}\left(\frac{1 + t^4}{1 - t^4}\right) \right. \\ &+ \left. \sum_{j=1}^{\infty} \frac{1}{\Gamma(jN - h + \frac{1}{2})} P^{jN-h}_{-\frac{1}{2}}\left(\frac{1 + t^4}{1 - t^4}\right) \right]. \end{aligned} \tag{9}$$

The first term can be transformed, by means of the identity[3]

$$P_{-\frac{1}{2}}(\cosh\eta) = [2/\pi \cosh(\tfrac{1}{2}\eta)] K(\tanh \tfrac{1}{2}\eta), \tag{10}$$

to

$$\tfrac{1}{2}a_0 = (2/\pi) K(t^4) \tag{11}$$

where $K(x)$ is a complete elliptic integral of the first kind. Other terms in the series can be written as linear combinations of complete elliptic integrals. However, they are usually negligible since for large N the asymptotic estimate[3]

$$\begin{aligned} &P^{jN+h}_{-\frac{1}{2}}\left(\frac{1 + t^4}{1 - t^4}\right) \\ &= \frac{1}{(jN + h)!} t^{2(jN+h)} \left[1 + O\left(\frac{1}{jN + h + 1}\right) \right] \end{aligned} \tag{12}$$

is valid.

In many cases the summations indicated in Eq. (1)

[2] A. A. Maradudin, E. W. Montroll, and G. H. Weiss, *Lattice Dynamics in the Harmonic Approximation*, (Academic Press Inc., New York, 1963).

[3] A. Erdelyi, et al., *Higher Transcendental Functions* (McGraw-Hill Book Company, Inc., New York, 1953), Vol. *II*.

can be carried out in closed form. Consider the one-dimensional Green's function

$$g(n;\cosh\varphi)=\frac{1}{N}\sum_{j=1}^{N}\frac{\cos(2\pi jn/N)}{\cosh\varphi+\cos(2\pi j/N)}. \tag{13}$$

The Fourier coefficients are

$$a_j=2(-1)^j e^{-j|\varphi|}/\sinh|\varphi|. \tag{14}$$

Hence, after some algebra we find

$$g(n,\cosh\varphi)=\frac{(-1)^n}{\sinh|\varphi|}\frac{\cosh(\frac{1}{2}N-n)|\varphi|}{\sinh\frac{1}{2}N|\varphi|} \tag{15}$$

for N even, and

$$g(n,\cosh\varphi)=\frac{(-1)^n}{\sinh|\varphi|}\frac{\sinh(\frac{1}{2}N-n)|\varphi|}{\sinh\frac{1}{2}N|\varphi|} \tag{16}$$

for N odd. Any sum which can be put into the form

$$g(n;\cosh\varphi_1,\cosh\varphi_2,\cdots\cosh\varphi_k)=\frac{1}{N}\sum_{j=1}^{N}\frac{\cos(2\pi jn/N)}{\prod_{m=1}^{k}[\cosh\varphi_m+\cos(2\pi j/N)]} \tag{17}$$

can be evaluated in a similar manner by expanding the demoninator into partial fractions. The sum

$$g(n;\cos\varphi)=\frac{1}{N}\sum_{j=1}^{N}\frac{\cos(2\pi jn/N)}{\cos\varphi+\cos(2\pi j/N)} \tag{18}$$

can be evaluated by a similar technique (provided that φ is chosen so that the denominator doesn't vanish) by considering the function $g(n;\cos\varphi+i\epsilon)$. The Fourier coefficients of the latter function are defined, in contrast to those of the function in Eq. (18). In the final result, the limit $\epsilon\to 0$ is to be taken; all of this is equivalent to taking the principal values of the Fourier coefficients.

An interesting example is provided by an expression derived by Mazur and Montroll for the momentum autocorrelation function for a mass in a monotonic lattice.[4] This expression is

$$\rho(t)=\frac{1}{2N+1}\sum_{j=-N}^{N}\cos\left(\omega_L t\sin\frac{\pi j}{2N+1}\right). \tag{19}$$

By a slight modification of our technique, and by using the formula

$$\cos(x\sin\theta)=J_0(x)+2\sum_{n=1}^{\infty}J_n(x)\cos n\theta, \tag{20}$$

we find that Eq. (19) can be transformed to

$$\rho(t)=J_0(\omega_L t)+\frac{1}{N+\frac{1}{2}}\sum_{n=1}^{\infty}\left\{1+\frac{2\cos[\frac{1}{2}n\pi(N+1)/(2N+1)]\sin[Nn\pi/(4N+2)]}{\sin[n\pi/(4N+2)]}\right\}J_n(\omega_L t). \tag{21}$$

In the limit of $N\to\infty$, $\rho(t)$ is approximated by $J_0(\omega_L t)$ which tends to zero as $t\to\infty$. It is clear, however, from Eq. (19) that $\rho(t)$ is an almost periodic function. An idea of the approximation involved can be gained by examining the sum which appears in Eq. (21). The coefficients of $J_n(\omega_L t)$ vanish with $(2N+1)^{-1}$ for all values of n except for those in which the sine term in the denominator vanishes. For this case we must have

$$n\pi/(4N+2)=k\pi,\qquad k=1,2,\cdots, \tag{22}$$

and the term in brackets becomes $2N+1$. Therefore, excepting terms which are $O(N^{-1})$, we have the approximation

$$\rho(t)=J_0(\omega_L t)+2\sum_{j=1}^{\infty}J_{2(2N+1)j}(\omega_L t). \tag{23}$$

Thus we see that the approximation $\rho(t)=J_0(\omega_L t)$ will generally be valid for $\omega_L t\ll 2(2N+1)$; otherwise, other terms from the series become significant. These results can also be obtained by physical arguments, but the present discussion provides a quantitative estimate of deviations from the integral approximation.

The extension of the present method to higher dimensions is straightforward. Suppose that the sum of interest is

$$S_N(r_1,r_2,\cdots r_k)=\sum_{j_1=1}^{N}\cdots\sum_{j_k=1}^{N}F\left(\cos\frac{2\pi j_1}{N},\cos\frac{2\pi j_2}{N},\cdots,\cos\frac{2\pi j_k}{N}\right)\times\cos\frac{2\pi j_1 r_1}{N}\cdots\cos\frac{2\pi j_k r_k}{N}, \tag{24}$$

and $F(\cos\theta_1,\cdots\cos\theta_k)$ has the Fourier representation

$$F(\cos\theta_1,\cdots,\cos\theta_k)=\sum_{n_1=0}^{\infty}\cdots\sum_{n_k=0}^{\infty}A_{n_1,\cdots n_k}\times\cos n_1\theta_1\cos n_2\theta_2\cdots\cos n_k\theta_k, \tag{25}$$

where

$$A_{n_1,n_2,\cdots,n_k}=\frac{2^p}{\pi^k}\int_0^{2\pi}\cdots\int F(\cos\theta_1,\cos\theta_2,\cdots\cos\theta_k)\times\cos n_1\theta_1\cdots\cos n_k\theta_k\,d\theta_1\cdots d\theta_k, \tag{26}$$

[4] P. Mazur and E. W. Montroll, J. Math. Phys. 1, 70 (1960).

and p is the number of indices $(n_1, \cdots n_k)$ which are equal to zero. Then a rearrangement of summation's yields an alternate expression for $S_N(r_1, \cdots r_k)$ as

$$S_N(r_1, r_2, \cdots r_k) = \frac{1}{2^k} \sum_{j_1=-\infty}^{\infty} \cdots \sum_{j_k=-\infty}^{\infty} A_{|j_1N+r_1|,|j_2N+r_2|,\cdots|j_kN+r_k|}. \quad (27)$$

Using these results we see that the finite sum of Eq. (3) can be transformed into an infinite sum involving the integrals

$$A_{n_1,n_2,n_3} = \frac{1}{\pi^3} \times \iiint_0^{2\pi} \frac{\cos n_1\theta_1 \cos n_2\theta_2 \cos n_3\theta_3 \, d\theta_1 \, d\theta_2 \, d\theta_3}{z + \alpha_1 \cos\theta_1 + \alpha_2 \cos\theta_2 + \alpha_3 \cos\theta_3}, \quad (28)$$

$$|z| > |\alpha_1 + \alpha_2 + \alpha_3| \quad .$$

Tables of these integrals are available for the case $\alpha_1 = \alpha_2$, $\alpha_3/\alpha_1 = 1$, 2, 4, 8, and 16, and $z/\alpha_2 = (2 + \alpha_3/\alpha_2)\beta$, where $\beta^{-1} = 0.0\ (.01)\ 1.00$.[5] The first approximation to the sum in Eq. (3) is just the value of the integral approximation to the Green's function. Further correction terms can be derived from larger N by using the asymptotic evaluation[5]

$$\frac{1}{\pi^3} \iiint_0^{2\pi} \frac{\cos n_1\theta_1 \cos n_2\theta_2 \cos n_3\theta_3 \, d\theta_1 \, d\theta_2 \, d\theta_3}{z + \alpha_1 \cos\theta_1 + \alpha_1 \cos\theta_2 + \alpha_3 \cos\theta_3} \sim 4\left(\frac{\alpha_3}{\alpha_1}\right)^{\frac{1}{2}} \frac{e^{-AR}}{R}, \quad (29)$$

where

$$A^2 = 2(2 + \alpha_3/\alpha_1)[z/(2\alpha_1 + \alpha_3) - 1], \qquad R^2 = [n_1^2 + n_2^2 + (\alpha_1/\alpha_3)n_3^2]. \quad (30)$$

We therefore see that corrections to the integral approximation to the three-dimensional discrete Green's function fall off better than exponentially with increasing N. Thus an approximate expression for $g(n_1, n_2, n_3; z)$ is

$$g(n_1, n_2, n_3; z) \sim \tfrac{1}{8}(A_{n_1,n_2,n_3} + A_{N-n_1,n_2,n_3} + A_{n_1,N-n_2,n_3} + A_{n_1,n_2,N-n_3} + A_{n_1,N-n_2,N-n_3} + A_{N-n_1,n_2,N-n_3} + A_{N-n_1,N-n_2,n_3} + A_{N-n_1,N-n_2,N-n_3}). \quad (31)$$

[5] A. A. Maradudin, E. W. Montroll, G. H. Weiss, R. Herman, and H. W. Milnes, Mem. Acad. Roy. Belg. 14, 7 (1960).

Generally only one of the A's need be evaluated as a triple integral, and the remaining terms can be approximated as in Eq. (29).

Finally, we mention some generalizations. The present method is clearly not restricted to functions of the form $F(\cos\theta)$, but any summation of the form

$$S_N = \sum_{j=1}^{N} F\left(\frac{2\pi j}{N}\right) \quad (32)$$

can be transformed to an infinite, but sometimes more rapidly convergent, series by calculating the Fourier coefficients of $F(\theta)$ and using Eq. (6). Also, anite products of the form

$$P_N = \prod_{j=1}^{N} F\left(\frac{2\pi j}{N}\right) \quad (33)$$

can be calculated by taking logarithms, and thereafter following the steps just outlined. Further, more general forms of Poisson summation can be employed to generate identies such as those we have employed for Fourier-series representations. As an example, let $\{\varphi_n(x)\}$ be an orthonormal sequence of functions, and let $F(x)$ be expressible as

$$F(x) = \sum_{n=0}^{\infty} a_n\varphi_n(x). \quad (34)$$

If we assume that the a_n are such as to permit free interchanges of orders of summation, and that the sums

$$U_n(N) = \sum_{j=1}^{N} \varphi_n\left(\frac{2\pi j}{N}\right) \quad (35)$$

are known, then we can formally write

$$\sum_{j=1}^{N} F\left(\frac{2\pi j}{N}\right) = \sum_{n=0}^{\infty} a_n U_n(N). \quad (36)$$

An almost trivial example of this is provided by Tschebycheff polynomials. If $F(x)$ can be expanded as

$$F(x) = \sum_{n=0}^{\infty} a_n T_n(x), \quad (37)$$

then

$$\frac{1}{N} \sum_{j=1}^{N} F\left(\cos\frac{2\pi j}{N}\right) = \frac{1}{N} \sum_{j=1}^{N} \sum_{n=0}^{\infty} a_n T_m\left(\cos\frac{2\pi j}{N}\right) = \frac{1}{N} \sum_{n=0}^{\infty} a_n \sum_{j=1}^{N} \cos\frac{2\pi jn}{N} = \sum_{j=0}^{\infty} a_{jN}. \quad (38)$$

No detailed calculations have been undertaken with this formula.